



```

RESULT 2
VDP_HUMAN STANDARD: PRT: 962 AA.
ID VDP_HUMAN 060763:
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE General vesicular transport factor p115 (Transcytosis associated
  protein) (TAP) (Vesicle docking protein).
GN VDP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PHOSPHORYLATION SITE SRR-942.
RA MEDLINE=98148093; PubMed=9478999;
RX Soida M., Misumi Y., Yano A., Takami N., Ikehara Y.;
RT "Phosphorylation of the vesicle docking protein p115 regulates its
  association with the Golgi membrane.";
RL J. Biol. Chem. 273:5385-5388(1998).
CC -I- FUNCTION: GENERAL VESICULAR TRANSPORT FACTOR REQUIRED FOR
  INTERCISTERNAL TRANSPORT IN THE GOLGI STACK. IT IS REQUIRED FOR
  TRANSCYTOTIC FUSION AND/OR SUBSEQUENT BINDING OF THE VESICLES TO
  THE TARGET MEMBRANE. MAY WELL ACT AS A VESICULAR ANCHOR BY
  INTERACTING WITH THE TARGET MEMBRANE AND HOLDING THE VESICULAR
  AND TARGET MEMBRANES IN PROXIMITY (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN WHICH RECYCLES
  BETWEEN THE CYTOSOL AND THE GOLGI APPARATUS DURING INTERPHASE.
CC -I- DOMAIN: COMPOSED OF A GLOBULAR HEAD, AN ELONGATED TAIL (COILED-
  COIL). AND A HIGHLY ACIDIC C-TERMINAL DOMAIN.
CC -I- PTM: PHOSPHORYLATED IN A CELL CYCLE-SPECIFIC MANNER:
  PHOSPHORYLATED IN INTERPHASE BUT NOT IN MITOTIC CELLS.
CC PHOSPHORYLATED PROTEIN ASSOCIATES WITH THE GOLGI MEMBRANE;
CC PHOSPHORYLATION PROMOTES DISSOCIATION.
CC -I- SIMILARITY: BELONGS TO THE VDP/USOL/YBL047C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
DR EMBL: D86326; BAA25300.1; -.
DR MIM: 603344; -.
DR InterPro: IPR000225; Armadillo.
DR PROSITE: PS50176; ARM_REPEAT; UNKNOWN_1.
KW Transport; Protein transport; Golgi stack; Membrane; Coiled coil;
  KW Phosphorylation.
FT DOMAIN 1 637 GLOBULAR HEAD.
FT DOMAIN 638 930 COILED COIL (POTENTIAL).
FT DOMAIN 935 962 ASP/GLU-RICH (ACIDIC).
FT MOD_RES 942 942 PHOSPHORYLATION.
FT MUTAGEN 942 942 S->A: LOSS OF PHOSPHORYLATION.
SQ SEQUENCE 962 AA: 107906 MW: 2E748F2C1BC2B942 CRC64;

```

```

Query Match 15.8%; Score 168; DB 1; Length 962;
Best Local Similarity 20.2%; Pred. No. 0.018;
Matches 65; Conservative 62; Mismatches 85; Indels 110; Gaps 7;

```

```

QY 3 IINKLFLPDLAEEENVLDREFLNELDNVRAQLSOKDEKRDQVYITLDRDLEERNAT 62
DB 608 IFDEHFKFLVLEBQVYIKAYIKSEEDKKEEYKKTLEQHDN--IVHYKMKITREQDLQ 665
QY 63 VVSLQOALG---KAEMCSTLKKOMKYLEQOODE-----TKQAQ-- 98
DB 666 LEELRQOVSTLKQONEQLOLAVTQOVSOIQHKDQYINLKITQLGKDNQHGSSYSEGAOMN 725

```

```

QY 99 ----EAGRLRSKMTQMEIQLLOSQLEVEEMIRDMGVGQSA----- 138
DB 726 GIQPEIRGLRREIREIELKRNQELLOSQLETKEDSMIENKSSQTSQTNBOSSAIVAROSE 785
QY 139 -VEQLAVYCVSLK----- 150
DB 786 QVAELKQELATLTKLSQNSQVEITKLQTEKQELLOKTEAFKSVQGETTITAFKTTD 845
QY 151 -----KEYENLKEARKASGEVADKLKRDLFSSRSKLQTVSELDQAKLELKSQKD 201
DB 846 VEGRLSALLQETKELKNEIKALSEERTAIKEQLDSSNTIAIQTETKOKLELEITDSKE 905
QY 202 -----LQSADEKTEISLKKRL 216
DB 906 QDDLVLVLADQDKITLSLKNKL 927

```

```

RESULT 3
MYS2_DICD1 STANDARD: PRT: 2116 AA.
ID MYS2_DICD1
AC P08799;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin II heavy chain, non muscle.
GN MHCA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87092266; PubMed=3540939;
RA Warrick H.M., de lozanne A., Leitwand L.A., Spudis J.A.;
RT "Conserved protein domains in a myosin heavy chain gene from
  Dictyostelium discoideum.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
RN [2]
RP PHOSPHORYLATION SITES, AND MUTAGENESIS.
RC STRAIN=AX2;
RX MEDLINE=90353583; PubMed=2387408;
RA Lueck-Vietmeier D., Schleicher M., Grabatin B., Wipplier J.,
  Gerisch G.;
RT "Replacement of threonine residues by serine and alanine in a
  phosphorylatable heavy chain fragment of Dictyostelium myosin II.";
RL FEBS Lett. 269:239-243(1990).
RN [3]
RP PHOSPHORYLATION SITES.
RX MEDLINE=88112226; PubMed=2828113;
RA Magle G., Noegel A., Scheel J., Gerisch G.;
RT "Phosphorylation of threonine residues on cloned fragments of the
  Dictyostelium myosin heavy chain.";
RL FEBS Lett. 227:71-75(1988).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
RX MEDLINE=95345066; PubMed=7619795;
RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
  Rayment I.;
RT "X-ray structures of the myosin motor domain of Dictyostelium
  discoideum complexed with MgADP. Bepx and MgADP.ALf4-";
RL Biochemistry 34:8960-8972(1995).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
RX MEDLINE=95345067; PubMed=7619796;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the
  truncated head of Dictyostelium discoideum myosin to 2.7-A
  resolution.";
RL Biochemistry 34:8973-8981(1995).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX MEDLINE=96206189; PubMed=8611530;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II).ADP.vanadate complex of the

```

RT Dictyostelium discoidium myosin motor domain to 1.9-A resolution.";  
RL Biochemistry 35:5404-5417(1996).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.  
RX MEDLINE=97452580; PubMed=9305951;  
RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;  
RT "X-ray structures of the MgAD, MgADgammaas, and MgAMPNP complexes  
RT of the Dictyostelium discoidium myosin motor domain.";  
RL Biochemistry 36:11619-11628(1997).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.  
RX MEDLINE=98070605; PubMed=9405148;  
RA Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;  
RT "X-ray crystal structure and solution fluorescence characterization  
RT of Mg.2'(3')-O-(N-methylanthraniloyl) nucleotides bound to the  
RT Dictyostelium discoidium myosin motor domain.";  
RL J. Mol. Biol. 274:394-407(1997).  
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE  
CC ACTIVITY THAT IS ACTIVATED BY ACTIN.  
CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES  
CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI  
CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS  
CC (MLC-2').  
CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL  
CC CORTEX.  
CC -1- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
CC MEROMYOSIN (LMN) AND 1 HEAVY MEROMYOSIN (HNM). IT CAN BE FURTHER  
CC SPLIT INTO 2 GLOBULAR SUBSEGMENTS (S1) AND 1 ROD-SHAPED  
CC SUBSEGMENT (S2).  
CC -1- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
CC -1- PPM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES  
CC THE ACTIN-ACTIVATED ATPASE ACTIVITY.  
CC -1- MISCELLANEOUS: DICTYOSTELIUM MYOSIN II HAS NO K(2)EDTA ATPASE  
CC ACTIVITY. PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1  
CC POSITION (668)  
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL, M14628; AAA33227.1; -.  
DR PIR; A26655; A26655.  
DR PIR; S00250; S00250.  
DR PDB; 1MMA; 03-DEC-97.  
DR PDB; 1MMD; 17-AUG-96.  
DR PDB; 1MMG; 03-DEC-97.  
DR PDB; 1MMN; 03-DEC-97.  
DR PDB; 1MND; 17-AUG-96.  
DR PDB; 1MNE; 17-AUG-96.  
DR PDB; 1VOM; 23-DEC-96.  
DR PDB; 1LVK; 28-JAN-98.  
DR Dictydb; DD01008; mhca.  
DR InterPro; IPR000048; IQ.  
DR InterPro; IPR004009; Myosin\_N.  
DR InterPro; IPR001609; myosin\_head.  
DR Pfam; PF00612; IQ; 2.  
DR Pfam; PF00063; myosin\_head; 1.  
DR Pfam; PF02736; myosin\_N; 1.  
DR PRINTS; PR00193; MYOSINHEAVY.  
DR ProDom; PD000355; myosin\_head; 1.  
DR SMART; SM00015; IQ; 1.  
DR SMART; SM00242; MYSC; 1.  
DR PROSITE; PS50096; IQ; 1.  
KW Myosin; Coiled coil; Actin-binding; ATP-binding; 3D-structure;  
KW Calmodulin-binding; Methylation; Alkylation; Phosphorylation.

```

FT DOMAIN 1 761 MYOSIN HEAD-LIKE.
FT DOMAIN 762 791 IQ.
FT DOMAIN 817 2116 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP.
FT DOMAIN 638 660 ACTIN-BINDING.
FT DOMAIN 738 752 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (DI-) (POTENTIAL).
FT MOD_RES 678 678 ALKYLATION (SH-1).
FT MOD_RES 1823 1823 PHOSPHORYLATION (BY HCK).
FT MOD_RES 1833 1833 PHOSPHORYLATION (BY HCK).
FT MOD_RES 2029 2029 PHOSPHORYLATION (BY HCK).
SQ SEQUENCE 2116 AA; 243871 MW; 2FC3770BB1EE56A1 CRC64;

Query Match 15.4%, Score 164.5; DB 1; Length 2116;
Best Local Similarity 25.4%; Pred. No. 0.063;
Matches 68; Conservative 56; Mismatches 87; Indels 57; Gaps 11;

OY 5 NKLFFDLAEEENVID-REFLNELDNRAQLSQND---KERDSQYIITDRLTLEERN 60
   :||| | |||| | | | | | | | | | | | | | | | | | | | | | | | |
DB 852 DKLESLKDTESNVIDLQROLKAERETLKAMYSDALEAQAQRELEINVEHSELEDERK 911
   :||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 61 ATTVSLQ-----QALKAEMLCSTLKKKKYLYEQQOODEFQAOEAGRLRS 106
   :||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 912 LALENDLQNKSVKRYVDLEPELOEEOKLRTLEKLKKKYEELPEEKRYVDGSDPTIS 971
   :||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 107 KMKTEQTELLLSQLPEVEEMI---RDMGV-----GQSAVOLAVYCVS----- 148
   :||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 972 R---LEKIKDELQKVEELTESFSESKDKLEKTRVRLGSELDLIVRLDSETRKSE 1028
   :||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 149 ----LKKEYENLKERKA-SCGVADKLKRLDFSSSKSLQTYSELDO-----AKLEIK 196
   :||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1029 LIRQKKKLEELKQVQELAETAAKLAQE--AAKKKLGQETTELNEKFNSEVTAARHVE 1086
   :||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 197 SAQKDLS---ADKEIMSLKKKLLMLQ 220
   :||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1087 KSKKTLSEQLAVVNNLEDEKKKNRDALE 1114
   :||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
TANA_XENLA STANDARD: PRT: 1744 AA.
ID TANA_XENLA
AC Q01550;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tanaabin.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Tadpole head;
RX MEDLINE=92398961; PubMed=1524825;
RA Hemmati-Brivanlou A., Mann R.W., Harland R.M.;
RT "A protein expressed in the growth cones of embryonic vertebrate
   neurons defines a new class of intermediate filament protein.";
RL Neuron 9:417-428(1992).
CC -1- TISSUE SPECIFICITY: GROWTH CONES OF EMBRYONIC VERTEBRATE NEURONS.
CC -1- DEVELOPMENTAL STAGE: IS EXPRESSED IN THE NEURDIA AND PERSIST
   DURING EMBRYOGENESIS IN THE BRAIN, CRANIAL NERVES, AND SPINAL
   CORD.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC -----
DR EMBL: M9387; AAA49966.1; -
DR PIR: JH0720; JH0720.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; filament. 2.
DR PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Neurone.
FT DOMAIN 1 12 HEAD.
FT DOMAIN 13 314 ROD.
FT DOMAIN 315 1744 TAIL.
FT DOMAIN 8 48 COIL 1A.
FT DOMAIN 49 60 LINKER 1.
FT DOMAIN 61 156 COIL 1B.
FT DOMAIN 157 179 LINKER 12.
FT DOMAIN 180 193 COIL 2A.
FT DOMAIN 194 199 LINKER 2.
FT DOMAIN 200 314 COIL 2B.
SQ SEQUENCE 1744 AA: 199561 MW: 6502EAC9FE6C4E93 CRC64:

```

```

Query Match 15.3% Score 163.5; DB 1; Length 1744;
Best Local Similarity 25.8%; Pred. No. 0.058;
Matches 60; Conservative 61; Mismatches 73; Indels 39; Gaps 9;

```

```

QY 12 AOEEENVLDREFLNEDNVRALQSQDKERDSQVITDRLDTEERNATVSLQALG 71
   |||||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 33 ALEREN-----ELRKREIHSLSKSEKCMKKHHEMM-KLRDALDDGHEMV----- 80
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 72 KAENLCSTLKQMKYLEQODETKAOEAEAGRLSKKMTQIETLLQSOL-----P 123
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 81 QAEWVRDSTVEIEFVQKRCLEKQARDAKELSESKLLEETRAIWKERLQLEA 140
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 124 EEEEMIRMGVQSAEVLAVYCSLKKEYNLEKARKA--SGEYADLRK--DLFSS-- 177
   |||||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 141 ELEILNRHEEKALMEB---EIASFSQRLNFRVAPAFKPEVDVQARLSEIWCQAV 197
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 178 ---RSKLTQVYSELDAKLELKSQKD-----LQSAKDEIMSLKKRLTMLQ 220
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 198 EEYKSEVSVLEAGLSSEKKNLRKYLEENKONRLLOSIDKELVSLKMKREALE 250
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```

```

RESULT 5
G160_MOUSE STANDARD; PRT; 1325 AA.
ID G160_MOUSE
AC P55937;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Golgin-160 (Male-enhanced antigen-2) (MEA-2).
GN GOLGA3 OR MEA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Testis;
RX MEDLINE=97217683; PubMed=9063644;
RA Kondo M., Sutoh S.;
RT "Cloning and molecular characterization of cDNA encoding a mouse
RT male-enhanced antigen-2 (Mea-2): a putative family of the golgi
RT autoantigen."
RL DNA Seq. 7:71-82(1997).
CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN SPERMATOGENESIS AND/OR
CC TESTIS DEVELOPMENT. PROBABLY IDENTICAL WITH THE SEROLOGICALLY
CC DETECTABLE MALE ANTIGEN (SDM).
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS. TRANSCRIPTS CAN BE
CC FOUND IN SPERMATIDS DURING SPERMATOGENESIS. NO EXPRESSION IN
CC LEYDIG CELLS, SPERMATOGENIA, OR SPERMATOCYTES.
CC -1- SIMILARITY: HIGH, TO HUMAN GOLGIN-160
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1, MET-19 OR MET-30 IS THE
CC INITIATOR.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: D78270; BAA19612.1; -
DR HSSP: P18852; ISCG.
DR MGD; MGI:96958; Golga3.
KW Spermatogenesis; Developmental protein.
FT DOMAIN 201 204 POLY-ALA.
SQ SEQUENCE 1325 AA: 149880 MW: 323063696ZC68780 CRC64:

```

```

Query Match 15.1% Score 161; DB 1; Length 1325;
Best Local Similarity 23.1%; Pred. No. 0.06;
Matches 57; Conservative 48; Mismatches 90; Indels 52; Gaps 7;

```

```

QY 13 QEEENVLDREFLNEDNVRALQSQDKERDSQVITDRLDTEERNATVSL 66
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 966 REHNSITETALAKREADVQLNLOVAVLQKREEDKQKOLVQALQSLKEMEVNSL 1025
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 67 QOALGKA-----EMLCSTLKQMKYLEQO-----ODET 94
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 1026 KEQMAAARIEAGNRRHFKATLELSEYKKELQAKENHYVQLQAEVDELQIDQDKHSQEI 1085
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 95 KOAEAEAGRLSKKMTQIETLLQSOLPEVEEMIRMGVQSAVEQALVYCSLKKEYE 154
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 1086 AQFQTELEARTQLQ-----LQKRLK--DEQSQQPGSQQMEDLQKWLQDKREBIQ 1135
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 155 NLKEPARKSGVADLRKLDLFSSRSKLTQVYSELDAKLELKSQKDLQSAKDEIMSLKK 214
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 1136 SLKQDLDTEQG---KKELEGTQDTIQTIKSELEMYQEDLSFTQKXKFMQAKVSELKN 1192
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 215 KL-TMLQ 220
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 1193 NMTLLQ 1199
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```

```

RESULT 6
MYS_AEOIR STANDARD; PRT; 1938 AA.
ID MYS_AEOIR
AC P24733;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, striated muscle.
DE Aequipecten irradians (Bay scallop).
OS Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoida;
OC Pectinoidea; Pectinidae; Argopecten.
OX NCBI_TaxID=31199;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adductor muscle;
RX MEDLINE=92011595; PubMed=1917970;
RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;
RT "Complete primary structure of a scallop striated muscle myosin heavy
RT chain. Sequence comparison with other heavy chains reveals regions
RT that might be critical for regulation."
RL J. Biol. Chem. 266:18469-18476(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Adductor muscle;
RX MEDLINE=91088319; PubMed=2263488;
RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;
RT "Nucleotide sequence of full length cDNA for a scallop striated
RT muscle myosin heavy chain."
RL Nucleic Acids Res. 18:7158-7158(1990).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 777-836.
RX MEDLINE=94173332; PubMed=8127365;

```

OY	66	LOCALGRAMEMICSTLTKQMKYLEGQODETRKQAQBAGRLSRKMTMQIELLOSQLEPV	125
Db	1115	LEELEIEARNARAVEVQRALNRELELEGRLDEAG-----ATSAELINKRRE-AEL	1168
OY	126	BEMIRDMGVGSAAVEQLLVVCYSLKKEEENIKRKARKSGAEVADKLRKDLFSSRSKLOTVY	185
Db	1169	LKIRLD---EASLSQHNAQSIALRRKHQD-----ANENADQV-----DOLQKVK	1211
OY	186	SELDOAKLELKSACKDIQS	204
Db	1212	SLEKEDKKDLKREMDLES	1230
<hr/>			
RESULT	7		
ID	USOI_YEAST	STANDARD:	PRT: 1790 AA.
AC	P25386;		
DT	01-MAY-1992 (Rel. 22, Last created)		
DT	01-MAY-1992 (Rel. 22, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
Dc	Intracellular protein transport protein USOI.		
GN	USOI OR INT1 OR YDL058W.		
OC	Saccharomyces cerevisiae (Baker's yeast).		
CC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
CC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.		
OX	NCBI_TaxID=4932;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-X2180-1A;		
RX	MEDLINE=91185402; PubMed=2010462;		
RA	Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,		
RA	Yamasaki M.;		
RT	"A cytoskeleton-related gene, usoi, is required for intracellular		
RT	protein transport in Saccharomyces cerevisiae.";		
RL	J. Cell Biol. 113:245-260(1991).		
RJ	[2]		
RP	SEQUENCE OF 782-1790 FROM N.A.		
RA	Hostetter M.K., Herman D.J., Bendel C.M., McCiellan M., Tao N.,		
RA	Kendrick K.E.;		
RN	Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.		
RP	[3]		
RP	SEQUENCE OF 1-8 FROM N.A.		
RA	Bai Y., Symington L.S.;		
RL	Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.		
CC	-I- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI		
CC	COMPLEX.		
CC	-I- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR		
CC	MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE		
CC	ER AND THE GOLGI COMPLEX.		
CC	-I- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED		
CC	OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL		
CC	COLLED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.		
CC	-I- SIMILARITY: BELONGS TO THE VDP/USOI/YBL047C FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way com-		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL: X54378; CAA38253.1;		
DR	EMBL: L03188; AAB00143.1;		
DR	EMBL: U53668; AAB66659.1;		
DR	PIR: A38455; A38455.		
DR	HSSP: P80220; IDIP.		
DR	SCD: S0002216; USOI.		
DR	InterPro: IPR002017; Spectrin.		
KW	Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.		
FT	DOMAIN	1 724	GLOBULAR HEAD.
FT		725 1790	COILED COIL (POTENTIAL).
FT	DOMAIN	465 487	CHANGED (HYPER-HYDROPHILIC).

```

FT DOMAIN 991 1790 DISPENSABLE FOR THE PROTEIN FUNCTION.
FT DOMAIN 1172 1786 ASP/GLU-RICH (ACIDIC).
FT CONFLICT 847 847 G -> E (IN REF. 2).
FT CONFLICT 924 924 E -> K (IN REF. 2).
FT CONFLICT 1253 1253 V -> I (IN REF. 2).
FT CONFLICT 1319 1319 I -> V (IN REF. 2).
FT CONFLICT 1461 1461 N -> S (IN REF. 2).
FT CONFLICT 1581 1581 G -> S (IN REF. 2).
FT CONFLICT 1600 1600 I -> V (IN REF. 2).
FT CONFLICT 1661 1661 R -> S (IN REF. 2).
FT CONFLICT 1772 1772 D -> DEEDDEE (IN REF. 2).
SO SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9FD4818 CRC64;

Query Match 14.9%; Score 159; DB 1; Length 1790;
Best Local Similarity 23.7%; Pred. No. 0.1;
Matches 61; Conservative 57; Mismatches 85; Indels 54; Gaps 9;

QY 1 RTINKLFFDLAQ---EEENVLDREFLK-----NELDVRAQLSOKDKERKDSQVI 48
DB 1369 RKLNEGSSITQEVSEKINFLDELRLQNNELKAKEIDNTRSEL-EKVSLSND---- 1423
QY 49 IDLTLDTEERRATVYSIQAL-----GKAEMLCSTLKKOMKYLEQODETQAOEE 100
DB 1424 -----ELLEEKONTIKSLQDELISYKDKITRNDKELISIERDNKRLDLESLKQALRAQES 1478
QY 101 AGRLSKMKTMBOJELLQSLPEVEEMIRDMVGOSAVE---OLAVYCVSLKKEVENL 156
DB 1479 KAYVEGLAKLEEESSKEKAELEKEMKKL---ESTITESMETLKSMTETKRSDEKL 1535
QY 157 KEARRASGEVADKLR-----KDLFSSRSKLQTVY---SELDAQLELKSQNO 199
DB 1536 EQSKSAEEDIKNLQHEKSDILSRINESEKDIIEELKSLRIEAKSGSELEYKQGLNNAQ 1595
QY 200 KDLOSADKEIMSLKKL 216
DB 1596 EKIRINAEENTVLSKL 1612

RESULT 8
MYHD_HUMAN STANDARD; PRT; 1938 AA.
ID MYHD_HUMAN 095252;
DC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, skeletal muscle, extraocular (MyHC-ec).
GN MYH13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Extraocular muscle;
RX MEDLINE=99318869; PubMed=10388558;
RA Weiss A., Schiaffino S., Leinwand L.A.;
RT "Comparative sequence analysis of the complete human sarcomeric myosin
RT heavy chain family: implications for functional diversity.";
RL J. Mol. Biol. 290:61-75(1999).
RN [2]
RP SEQUENCE OF 1917-1938 FROM N.A.
RC TISSUE-Extraocular muscle;
RX MEDLINE=99026150; PubMed=9806854;
RA Winters L.M., Briggs M.M., Schachar F.;
RT "The human extraocular muscle myosin heavy chain gene (MYH13) maps to
RT the cluster of fast and developmental myosin genes on chromosome 17.";
RL Genomics 54:188-189(1998).
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATOR LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.

```

```

CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTR: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LMN) AND 1 HEAVY MEROMYOSIN (HNM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF111782; AAD29948.1; -
DR EMBL: AF075248; AAC83241.1; -
DR HSSP: P08799; LMND.
DR MIM: 603487; -
DR InterPro: IPR000048; IQ.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 2.
DR Pfam: PF00663; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
KW Myosin: muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
KW MYOSIN HEAD-LIKE.
FT DOMAIN 1 784
FT MYHD_HUMAN 843 1938
FT NP_BIND 179 186
FT DOMAIN 659 681
FT MYHD_HUMAN 761 775
FT MOD_RES 130 130 METHYLATION (TRI-) (POTENTIAL).
FT MOD_RES 699 699 ALKYLATION (SH-1) (POTENTIAL).
FT MOD_RES 709 709 ALKYLATION (SH-2) (POTENTIAL).
SO SEQUENCE 1938 AA; 223678 MW; 1F6D006416381CD5 CRC64;

Query Match 14.9%; Score 159; DB 1; Length 1938;
Best Local Similarity 23.3%; Pred. No. 0.11;
Matches 57; Conservative 48; Mismatches 86; Indels 54; Gaps 8;

QY 13 QEEENVL-----DREFLNELDVRAQLSOKDKER-----ROSOVIIDLTLDTEERN 60
DB 1296 EKEESLISQLTKRSKQALFQOLEELKRWEEEEKANAMAHALQSRHCCDLIREYEEEQ 1355
QY 61 ATVVSIAQALGKAE-----MLCSTLKKOMKYLEQODETQAOEEAGRLR 105
DB 1356 EAKKELQALSKANSEVAMQWTKYETDAIQTELEAKKKLAQRLQAEKETA---N 1412
QY 106 SKMKTMBOJELLQSLPEVEEMIRDMVGOSAVEQLAVYCVSLKKEVENLK----- 157
DB 1413 SKCASLETKQRLQG---EVEDLMRDL-----ERSHTACATLDKKORNPDKVLAEWKQ 1462
QY 158 --EARKASGEVADK---LRKDLFSSRSKLQTVYSSELDAQLELKSAPKDIQSDAKKEIMS 211
DB 1463 KLDESQALEAQAQESRSLSTELFKRNAYEEVVDLETLRRENKNLOEETISDLTEQIAE 1522
QY 212 LKKKL 216

```

Db 1523 TGNL 1527

RESULT 9

MYH7\_RAT STANDARD; PRT; 1935 AA.

AC P02564;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Myosin heavy chain, cardiac muscle beta isoform (MyHC-beta).

GN MYH7.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN (1)

RP SEQUENCE FROM N.A.

RC TISSUE=Heart;

RX MEDLINE=90016823; PubMed=2798112;

RA Kraft R., Bravo-Zehnder M., Taylor D., Leinwand L.A.;

RT "Complete nucleotide sequence of full length cDNA for rat beta cardiac myosin heavy chain.";

RL Nucleic Acids Res. 17:7529-7530(1989).

RN (2)

RP DISCUSSION OF SEQUENCE.

RX MEDLINE=90133919; PubMed=2614840;

RA McNally E.M., Kraft R., Bravo-Zehnder M., Taylor D., Leinwand L.A.;

RT "Full-length rat alpha and beta cardiac myosin heavy chain sequences. Comparisons suggest a molecular basis for functional differences.";

RL J. Mol. Biol. 210:665-671(1989).

RN (3)

RP SEQUENCE OF 1524-1935 FROM N.A.

RX MEDLINE=62220036; PubMed=7045682;

RA Mahdavi V., Periasamy M., Nadal-Ginard B.;

RT "Molecular characterization of two myosin heavy chain genes expressed in the adult heart.";

RL Nature 297:659-664(1982).

RN (4)

RP SEQUENCE OF 1871-1935 FROM N.A.

RC STRAIN=Wistar; TISSUE=Heart;

RX MEDLINE=65179510; PubMed=6241892;

RA Mahdavi V., Lompre A.M., Chambers A.P., Nadal-Ginard B.;

RT "Cardiac myosin heavy chain isozymic transitions during development and under pathological conditions are regulated at the level of mRNA availability.";

RL Eur. Heart J. 5:181-191(1984).

RT FUNCTION: MUSCLE CONTRACTION.

CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).

CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.

CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.

CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.

CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).

CC -1- MISCELLANEOUS: THE CARDIAC ALPHA ISOFORM IS A 'FAST' ATPASE MYOSIN, WHILE THE BETA ISOFORM IS A 'SLOW' ATPASE.

CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch).

CC -----

CC EMBL: X15939; CAA34065.1; -

CC DR EMBL: J00752; AAA41654.1; -

CC DR EMBL: M32698; AAA41659.1; -

CC DR PIR: S06006; S06006.

CC DR PIR: A02989; A02989.

CC DR HSSP: P08799; 1MMD.

CC DR InterPro: IPR000048; IQ.

CC DR InterPro: IPR004009; Myosin\_N.

CC DR InterPro: IPR002928; Myosin\_Tail.

CC DR InterPro: IPR001609; myosin\_head.

CC DR Pfam: PF00612; IQ; 2.

CC DR Pfam: PF00063; myosin\_head; 1.

CC DR Pfam: PF02736; Myosin\_N; 1.

CC DR Pfam: PF01576; Myosin\_Tail; 1.

CC DR PRINTS: PR00193; MYOSINHEAVY.

CC DR PRODOM: PD000355; myosin\_head; 1.

CC DR SMART: SM00015; IQ; 1.

CC DR SMART: SM00242; MYSC; 1.

CC DR PROSITE: PS50096; IQ; 1.

CC KW Myosin; muscle protein; coiled coil; thick filament; Actin-binding; ATP-binding; Methylation; Alkylation; Multigene family;

CC KW Calmodulin-binding.

CC KM DOMAIN 1 780 MYOSIN HEAD-LIKE.

CC FT DOMAIN 781 810 IQ.

CC FT DOMAIN 840 1935 COILED COIL (POTENTIAL).

CC FT NP\_BIND 178 185 ATP.

CC FT DOMAIN 655 677 ACTIN-BINDING.

CC FT DOMAIN 757 771 ACTIN-BINDING.

CC FT MOD\_RES 129 129 METHYLATION (TRI-1) (POTENTIAL).

CC FT MOD\_RES 695 695 ALKYLATION (SH-1) (POTENTIAL).

CC FT MOD\_RES 705 705 ALKYLATION (SH-2) (POTENTIAL).

CC FT MOD\_RES 1529 1531 ALKYLATION (SH-2) (POTENTIAL).

CC FT CONFLICT 1731 1731 D -> H (IN REF. 3).

CC FT CONFLICT 1784 1784 N -> K (IN REF. 3).

CC FT CONFLICT 1851 1851 T -> N (IN REF. 3).

CC FT CONFLICT 1858 1858 R -> K (IN REF. 3).

CC FT SEQUENCE 1935 AA; 223082 MW; C8376C324A7BD82B CRC64;

CC -----

CC Query Match 14.9%; Score 158.5; DB 1; Length 1935;

CC Best Local Similarity 20.8%; Pred. No. 0.12;

CC Matches 62; Conservative 60; Mismatches 87; Indels 89; Gaps 9;

CC -----

CC QY 1 RTITNKLFDDAQBENVLDR-----EFLK----- 25

CC Db 1129 RAKVEKLRSDLSRELEISERLERAGATSVQIEMKKREAFOKMRDLERATLOHEAT 1188

CC QY 26 -----NELDNVRAQLSOKDEKRDSDQVITLNDLTLEERNATVVSLOQA 69

CC Db 1189 AALRRKHADSVALEGEIDVLRKQKLEKEKSEFKLELDVSNMQ-----IIRA 1241

CC QY 70 LGKAEMLCSTLKQMKYLEQOODETKQAQBEAGRLSKMKT-----MQEILLLOSQ 121

CC Db 1242 KANLEKMKRTLEDQNMHRSAEFTQSRVNDLTRRAALQVENEGLSNQDEKELI-SQ 1300

CC QY 122 LP-----EVEEMIRDMGVGSAVEQLAVCYSLKKEYENLKEARKASCEV 166

CC Db 1301 LTRGLTYTQOLEDLKRLQLEEVAKAKNALAHALASARHCDCLREQYEETEFAVAELQRY 1360

CC QY 167 ADKLKRLDFFSSRSKLOT---VYSRLDQAKLELSAQKDLDSAPKEIMSLKKKLTMLQ 220

CC Db 1361 LSKANSEVAQMRFTYETDAIQRTETLEERAKKL--AQR-LQDAEVAEVAVAKGSSLE 1415

CC -----

CC RESULT 10

CC MYSS\_CYPCA STANDARD; PRT; 1935 AA.

CC AC Q90339;

CC DT 16-OCT-2001 (Rel. 40, Created)

CC DT 16-OCT-2001 (Rel. 40, Last sequence update)

CC DT 16-OCT-2001 (Rel. 40, Last annotation update)



DE Myosin heavy chain, fast skeletal muscle.  
 OS *Cyprinus carpio* (Common carp).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 CC Cypriniformes; Cyprinidae; Cyprinus.  
 CX NCBI\_Taxid=7962;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fast muscle;  
 RX MEDLINE=97352533; PubMed=9208928;  
 RA Hirayama Y., Watabe S.;  
 RT "Structural differences in the crossbridge head of temperature-associated myosin subfragment-1 isoforms from carp fast skeletal muscle.";  
 RT Eur. J. Biochem. 246:380-387(1997).  
 RN [2]  
 RP SEQUENCE OF 981-1935 FROM N.A.  
 RC TISSUE=Fast muscle;  
 RX MEDLINE=97176447; PubMed=9023993;  
 RA Imai J., Hirayama Y., Kikuchi K., Kakinuma M., Watabe S.;  
 RT "cDNA cloning of myosin heavy chain isoforms from carp fast skeletal muscle and their gene expression associated with temperature acclimation.";  
 RT J. Exp. Biol. 200:27-34(1997).  
 RN [3]  
 RP SEQUENCE OF 1387-1528 FROM N.A.  
 RX MEDLINE=95194396; PubMed=7887920;  
 RA Watabe S., Imai J., Nakaya M., Hirayama Y., Okamoto Y., Masaki H., Uozumi T., Hirono I., Aoki T.;  
 RT "Temperature acclimation induces light meromyosin isoforms with different primary structures in carp fast skeletal muscle.";  
 RL Biochem. Biophys. Res. Commun. 208:118-125(1995).  
 CC -1- FUNCTION: MUSCLE CONTRACTION.  
 CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -1- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPEITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE SI DOMAIN ARE SELECTIVELY ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (SL) AND 1 ROD-SHAPED SUBFRAGMENT (S2).  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or send an email to a license@sib-sib.ch) or send an email to a license@sib-sib.ch).  
 CC -----  
 DR EMBL: D89992; BAA22069.1; -;  
 DR EMBL: D50476; BAA09069.1; -;  
 DR EMBL: D43700; BAA07802.1; -;  
 DR HSSP: P08799; IAMD.  
 DR InterPro: IPR000048; IQ.  
 DR InterPro: IPR004009; Myosin\_N.  
 DR InterPro: IPR002928; Myosin\_N.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00612; IQ; 1.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR Pfam: PF02736; Myosin\_N; 1.  
 DR Pfam: PF01576; Myosin\_tail; 1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR ProDom: PD000335; myosin\_head; 1.  
 DR SMART: SM00015; IQ; 1.  
 DR SMART: SM00242; MYSC; 1.

DR PROSITE: PS50096; IQ; 1.  
 KW ATP-binding; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KW ATP-binding; Methylation; Alkylation; Calmodulin-binding;  
 KW Multigene family.  
 KW DOMAIN 1 MYOSIN HEAD-LIKE.  
 FT DOMAIN 1 782 811 IQ.  
 FT DOMAIN 812 839 HINGE.  
 FT DOMAIN 840 1935 COILED COIL (POTENTIAL).  
 FT NP\_BIND 178 185 ATP (POTENTIAL).  
 FT DOMAIN 659 681 ACTIN-BINDING.  
 FT DOMAIN 761 775 ACTIN-BINDING.  
 FT MOD\_RES 129 129 METHYLATION (TR-1) (POTENTIAL).  
 FT MOD\_RES 699 699 ALKYLATION (SH-1).  
 FT MOD\_RES 709 709 ALKYLATION (SH-2).  
 FT SEQUENCE 1935 AA; 221599 MW; 9A1244B67D63C83B CRC64;  
 SO  
 Query Match 14.7%; Score 157; DB 1; Length 1935;  
 Best Local Similarity 22.4%; Pred. No. 0.14;  
 Matches 57; Conservative 51; Mismatches 87; Indels 60; Gaps 8;  
 Oy 15 EENVLDREF-----LKNELDVNRALSKDKKRNDSOYIITLRLTLER 59  
 Db 1180 EESTLQHEATAALRKQADSVAELGEQIDNLQRYKQLEKESYKMEIDLTSMNE-- 1237  
 Oy 60 NATVSLQALGKAEMLCSTLKQKRYLEQOODETKQAEAGRLSKMT-----M 111  
 Db 1238 -----AVAKKANLEKMRTELDLSFKTSDEVRQNDNMNRALQTENGFSROL 1292  
 Oy 112 EQIELLIQSOLP-----EVEEMIRMGVGSQAVEQLAVYCVSLKREYNL 156  
 Db 1293 EEKEALV-SQLTRGQAVTQOIEELKRHIEEVKKNALAHAVASRRHCDLREQYEE 1351  
 Oy 157 KEAR-----KASGEVAD---KIKRDLFSSRSKQLQVVEIQAKLEKSAQKDQSA 205  
 Db 1352 QEAKEILOGRMSKASVEAQMRTKETDIAIORTLEBAKKRLAQ---RLDDEESIEAV 1408  
 Oy 206 DKEIMSLKKLTMLQ 220  
 Db 1409 NSKASLEKTRQRLQ 1423  
 RESULT 11  
 VDP\_MOUSE STANDARD; PRT; 941 AA.  
 ID VDP\_MOUSE  
 AC 092120;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE General vesicular transport factor p115 (Transcytosis associated protein) (TAP) (Vesicle docking protein) (Fragment).  
 GN VDP.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129; TISSUE=Thymus;  
 RA Han S.-J.;  
 RT "Identification of mouse TAP (transcytosis associated protein/p115)."  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: GENERAL VESICULAR TRANSPORT FACTOR REQUIRED FOR INTERCISTERNAL TRANSPORT IN THE GOLGI STACK; IT IS REQUIRED FOR TRANSCYTOTIC FUSION AND/OR SUBSEQUENT BINDING OF THE VESICLES TO THE TARGET MEMBRANE. MAY WELL ACT AS A VESICULAR ANCHOR BY INTERACTING WITH THE TARGET MEMBRANE AND HOLDING THE VESICULAR AND TARGET MEMBRANS IN PROXIMITY (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN WHICH RECYCLES BETWEEN THE CYTOSOL AND THE GOLGI APPARATUS DURING INTERPHASE (BY SIMILARITY).  
 CC -1- DOMAIN: COMPOSED OF A GLOBULAR HEAD, AN ELONGATED TAIL (COILED-COIL) AND A HIGHLY ACIDIC C-TERMINAL DOMAIN.



CC -1- PTM: PHOSPHORYLATED IN A CELL CYCLE-SPECIFIC MANNER;  
CC PHOSPHORYLATED IN INTERPHASE BUT NOT IN MITOTIC CELLS.  
CC DEPHOSPHORYLATED PROTEIN ASSOCIATES WITH THE GOLDI MEMBRANE;  
CC PHOSPHORYLATION PROMOTES DISSOCIATION (BY SIMILARITY).  
CC  
CC -1- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AF096868; AAC72967.1; -  
CC MGD: MGI:1928095; VDP.  
CC InterPro: IPR000225; Armadillo.  
CC InterPro: IPR000219; Rhogef.  
CC SMART: SM00325; Rhogef. 1.  
CC PROSITE: PS50176; ARM\_REPEAT: UNKNOWN\_1.  
CC Transport: Protein transport; Golgi stack; Membrane; Coiled coil;  
CC Phosphorylation.  
CC Non\_Ter 1  
CC DOMAIN <1 619 GLOBULAR HEAD.  
CC FT DOMAIN 620 912 COILED COIL (POTENTIAL).  
CC FT DOMAIN 917 941 ASP/GLU-RICH (ACIDIC).  
CC MOD\_RES 922 922 PHOSPHORYLATION (BY SIMILARITY).  
SQ SEQUENCE 941 AA; 105152 MW; AB5C10855CD7E508 CRC64;

Query Match 14.6%; Score 155.5; DB 1; Length 941;  
Best Local Similarity 19.5%; Pred. No. 0.084;  
Matches 64; Conservative 52; Mismatches 81; Indels 131; Gaps 8;

QY 15 EENVLDREFLK--NELDNV--RAQLSQDKERKDSQVYIIDRLDRLERNATVYSLQAL 70  
DB 587 EYMFIDHFTFLVKLEGVITKALYKSSSEKKEEV-----KKTLQHDNIVTHYKMMI 641  
QY 71 GKAEMLCSTLKKOMKYLBOODE-----TKQAO----- 98  
DB 642 REQDLQLEELKQVSTLKQCNQLOLAVTQQAQSQIQHKKDYQNLKVLQKDNHIGSHG 701  
QY 99 -----EAGRLRSKMTMEQIELLLSQLPPEVEMIRDMGVGSA--VEQLAVYC 146  
DB 702 DGAQVNGIOPFELISRLREIEFLKSQLLOGQLAKEXDSLLENLKSQASGSEQASATC 761  
QY 147 ----- 146  
DB 762 PPDPDQVAELKQELTALKSQLCSQSLBITRLQTEKNCILLQRAETLAKSVPEGESEHVS 821  
QY 147 -----VSLKKEVENLKEARKASGEVADKLKDLFSRSKLTQVYVSLDQAKTEL 195  
DB 822 AAKTTDVGRUSALQETKELKNEIKALSEERTATQKQDSSNSTITAILQTEKDLDEV 881  
QY 196 KSAQKD-----LQSADKEIMSLKKTL 216  
DB 882 TDSKKEQDLDLVLLADQDKITLSKSL 909

RESULT 12  
XCP\_E\_XENIA  
ID XCP\_E\_XENIA STANDARD: PRT; 1203 AA.  
AC P50533:  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Chromosome assembly protein XCAP-E.  
GN XCAP-E.  
OS Xenopus laevis (African clawed frog).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
CC Xenopodinae; Xenopus.  
NCBI\_TaxID=8355;

RN [1]  
CC SEQUENCE FROM N.A.  
RX MEDLINE-95042742; PubMed-7954811;  
RA Hirano T., Mitchison T.J.;  
RT "A heterodimeric coiled-coil protein required for mitotic chromosome  
RL Cell 79:449-458(1994).  
CC -1- FUNCTION: REQUIRED FOR BOTH ASSEMBLY AND STRUCTURAL MAINTENANCE OF  
CC MITOTIC CHROMOSOMES.  
CC -1- SUBUNIT: ASSOCIATES WITH XCAP-C PROBABLY AS HETERODIMER.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR. DURING CHROMOSOME ASSEMBLY IN  
CC MITOTIC EXTRACTS, XCAP-C/E WAS RECRUITED TO THE CHROMATIN AND  
CC FORMED A DISCRETE INTERNAL STRUCTURE WITHIN ASSEMBLED CHROMOSOMES.  
CC -1- DOMAIN: CONSISTS OF TWO PUTATIVE CENTRAL COILED-COIL REGIONS  
CC FLANKED BY PUTATIVE GLOBULAR REGIONS AT THE N- AND C-TERMINUS.  
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: U13674; AAA64680.1; -  
CC HSPB: P07751; 1A73.  
CC InterPro: IPR003405; SMC\_C.  
CC InterPro: IPR003395; SMC\_N.  
CC Pfam: PF02483; SMC\_C; 1.  
CC Pfam: PF02463; SMC\_N; 1.  
CC Mitosis: ATP-binding; Coiled coil; Nuclear protein.  
CC NP\_BIND 32 39  
CC DOMAIN 172 356 ATP (POTENTIAL).  
CC FT DOMAIN 397 513 COILED COIL (POTENTIAL).  
CC FT DOMAIN 670 1032 COILED COIL (POTENTIAL).  
CC FT DOMAIN 1095 1121 ALA/ASP-RICH (DA-BOX).  
SQ SEQUENCE 1203 AA; 136341 MW; 04323DD0027DF309 CRC64;

Query Match 14.6%; Score 155.5; DB 1; Length 1203;  
Best Local Similarity 22.0%; Pred. No. 0.11;  
Matches 48; Conservative 54; Mismatches 87; Indels 29; Gaps 5;

QY 16 ENVLDRE-----EFLNELDNVRAQLSQDKERKDSQVYIIDRLDRLERNATVYSLQ 67  
DB 707 KNYERYQLKQOQEMKKEEAEILLQTKLQSSYHKKQEE--LDSLKQITESEETLKNTK 764  
QY 68 QALGKAEMLCSTLKKOMKYLBOQOD-ETKQAQEEAGRLRSKM---KTMQIELLLSQSL 122  
DB 765 EVQKKAEEKFVLEHKMNAEERERELKQAKLDTKAKKADASNKKKKEKQGEVDALV 824  
QY 123 PEVEEMIRDMGVGSAQVQLAVYCVSLKKEYENLKEARKASGEVADKLKDLFSRSKTLQ 182  
DB 825 LELFELKREQ-----TTYKQIETVDEAMKAVQEQDMSADEVSKKKEAVK 870  
QY 183 TVYSELQAKLEKSAQKDLSADKEIMSLKKTLMLQ 220  
DB 871 KAQDELAROKETIMGHDKETIKTSSEAGLRENNDLQ 908

RESULT 13  
MYH4\_RABIT  
ID MYH4\_RABIT STANDARD: PRT; 1938 AA.  
AC Q28641:  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Myosin heavy chain, skeletal muscle, juvenile.  
OS Oryctolagus cuniculus (Rabbit).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
NCBI\_TaxID=9986;

```

RN [1]
SEQUENCE FROM N.A.
RC STRAIN-NEW ZEALAND WHITE: TISSUE-Skeletal muscle;
RA Maeda K., Hoshinova E., Roesch-Kleinkauf A., Schuster H., Gasperik J.,
RT Wittinghofer A.;
RT "Isolation, sequencing of myosin heavy chain cDNA from rabbit
RT skeletal muscle and a novel cosynthesis of S-1 fragment with the
RT essential and regulatory light chains.";
RL Submitted (Aug-1995) to the EMBL/Genbank/DBJ databases.
CC -! FUNCTION: MUSCLE CONTRACTION.
CC -! SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -! SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -! DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -! MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATTER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -! SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -! SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32574; AAA74199.1; -.
DR HSSP: P08799; 1MMD.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ: 2.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRODOM: PD000355; myosin_head; 1.
DR SMART: SMO0015; IQ: 1.
DR SMART: SMO0242; MYSC: 1.
DR PROSITE: PS50096; IQ: 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multi-gene family.
FT DOMAIN 1 783 MYOSIN HEAD-LIKE.
FT DOMAIN 784 813 IQ.
FT NP_BIND 842 1938 COILED COIL (POTENTIAL).
FT DOMAIN 179 186 ATP (POTENTIAL).
FT DOMAIN 658 680 ACTIN-BINDING (BY SIMILARITY).
FT DOMAIN 760 774 ACTIN-BINDING (BY SIMILARITY).
FT MOD_RES 35 35 METHYLATION (MONO-) (BY SIMILARITY).
FT MOD_RES 130 130 METHYLATION (TRI-) (BY SIMILARITY).
FT MOD_RES 552 552 METHYLATION (TRI-) (BY SIMILARITY).
FT MOD_RES 756 756 METHYLATION (SH-1) (BY SIMILARITY).
FT MOD_RES 698 698 ALKYLATION (SH-2) (BY SIMILARITY).
FT MOD_RES 708 708 ALKYLATION (SH-2) (BY SIMILARITY).
SQ SEQUENCE 1938 AA; 223064 MW; DBA8A2EC5B182626 CRC64;

```

```

Query Match 14.5%; Score 155; DB 1; Length 1938;
Best Local Similarity 23.8%; Pred. No. 0.19;
Matches 54; Conservative 49; Mismatches 80; Indels 44; Gaps 7;

```

```

Oy 24 LKNEIDNRAQLSKDKKRDQVLIIDRLRLERNAIVSLQALKAEMLCSTLKKQ 83
Db 1206 LGEQIDNLRVAKKLEKSEIKMEIDDLASNM-----TVSKAKGLMEKMCRTLEDQ 1258

```

```

Oy 84 MKYLEQODETK-----QAQEDAGRL-----RSKMTMQEIEL 117
Db 1259 VSELKTKKEEHRRLNDLSAQRARQTESGERSROLDKSDLSVLSKGAQFQIQL 1318
Oy 118 LOSQLEPEVEMIRDMGVQSAVEQLAVYCVSIKKEYENLKEARKASGEVADKLKRDFFS 177
Db 1319 KR-----QLEEFIKASALAHALQSAHRHCDLLREQYEEQEKALQRAMSKANSEVQW 1374
Oy 178 RSKLQTF-----VSELDQAKLEKSNOKDQSDKREIMSLKKKLTMLQ 220
Db 1375 RTKYETDAIQRTLELEAKKL--AQR-LQDAEEHVEAVNAKCASTLE 1418

RESULT 14
MYH3_RAT
ID MYH3_RAT STANDARD: PRT: 1940 AA.
AC P12847.
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, fast skeletal muscle, embryonic.
GN MYH3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A. PubMed=3783701;
RX MEDLINE=87060988;
RA Streicher E.E., Streicher M.-A., Perriard J.C., Perlasamy M.,
RA Nadal-Ginard B.;
RT "Complete nucleotide and encoded amino acid sequence of a mammalian
RT myosin heavy chain gene. Evidence against Intron-dependent evolution
RT of the rod.";
RL J. Mol. Biol. 190:291-317(1986).
CC -! FUNCTION: MUSCLE CONTRACTION.
CC -! SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -! SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -! DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -! PFM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -! MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATTER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -! SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -! SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X04267; CAA27817.1; -.
DR PIR: A24922; A24922.
DR HSSP: P08799; 1MMD.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ: 2.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.

```

DR ProDom; PD000355; myosin\_head; 1.  
 DR SMART; SM00015; IQ; 1.  
 DR SMART; SM00242; MYSC; 1.  
 DR PROSITE; PS50096; IQ; 1.  
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;  
 KW Multigene family.  
 FT DOMAIN 1 781 MYOSIN HEAD-LIKE.  
 FT DOMAIN 782 811 IQ.  
 FT DOMAIN 840 1933 COILED COIL (POTENTIAL).  
 FT NP\_BIND 179 186 ATP (POTENTIAL).  
 FT DOMAIN 656 678 ACTIN-BINDING.  
 FT DOMAIN 758 772 ACTIN-BINDING.  
 FT MOD\_RES 130 130 METHYLATION (TRT-) (POTENTIAL).  
 FT MOD\_RES 696 696 ALKYLATION (SH-2).  
 FT MOD\_RES 706 706 ALKYLATION (SH-2).  
 SQ SEQUENCE 1940 AA; 223857 MW; B5D546A5965A696 CRC64;

Query Match 14.5%; Score 155; DB 1; Length 1940;  
 Best Local Similarity 23.8%; Pred. No. 0.19; Mismatches 83; Indels 36; Gaps 7;  
 Matches 53; Conservative 51;

QY 24 LKNELDNVRALQSLQKREKRDQVITLRLDTLEERNATVVSLOALGKAEMLCSTLRKQ 83  
 DB 1204 LAEQIDMLQKVKQKLEKESEFKLEIDLSISVSE-----SVSKSKANLEKICRTLEDQ 1256  
 QY 84 MKYLEQOODETKQA-----QEEAGRLRKSMTMEQIELL-----QSOLPEVEEM 128  
 DB 1257 LSEARGNEETQRLSELTTQKSRLOTEAGELSRQLEEKESIVSLSRSKQAFQOQIEEL 1316  
 QY 129 IRDWGVSQSAVEQLAVY-----CVSLKKEYENLKEARKASGVADLRKDLFSSSKL 181  
 DB 1317 KRLQEEENKAKNALAHALQSSRHCCDLLRQYEEQEGCKAELOALSKANSEVAQMKTKY 1376  
 QY 182 QT-----VYSELDAQLELKSQKDLQSAQDKELMSLKKRLTMDQ 220  
 DB 1377 ETDALQRTLEELKAKKRL--AQR-LQOSEQVEAVNAKCKASLE 1416

RESULT 15  
 CENF\_HUMAN STANDARD; PRT; 3210 AA.  
 ID CENF\_HUMAN P49454; Q13246; Q13171;  
 AC P49454; Q13246; Q13171;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE CENP-F kinetochore protein (Centromere protein F) (Mitosis) (AH antigen).  
 GN CENP-F.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast carcinoma;  
 RA MEDLINE=95348175; Pubmed=7542657;  
 RA Liao H., Winkler R.J., Mack G., Rattner J.B., Yen T.J.;  
 RT "CENP-F is a protein of the nuclear matrix that assembles onto kinetochores at late G2 and is rapidly degraded after mitosis.";  
 RL J. Cell Biol. 130:507-518(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=95379848; Pubmed=7651420;  
 RA Zhu X., Mancini M.A., Chang K.-H., Liu C.-Y., Chen C.-F., Shan B., Jones D., Yang-Feng T.L., Lee W.-H.;  
 RT "Characterization of a novel 350-kilodalton nuclear phosphoprotein that is specifically involved in mitotic-phase progression.";  
 RL Mol. Cell. Biol. 15:5017-5029(1995).  
 RN [3]  
 RP SEQUENCE OF 2194-3210 FROM N.A.  
 RA MEDLINE=95336446; Pubmed=7612011;

RA Li Q., Ke Y., Kapp J.A., Fertig N., Medsger T.A. Jr., Joshi H.C.;  
 RT "A novel cell-cycle-dependent 350-kDa nuclear protein: C-terminal domain sufficient for nuclear localization.";  
 RL Biochem. Biophys. Res. Commun. 212:220-228(1995).  
 RN [4]  
 RP CHARACTERIZATION.  
 RA MEDLINE=95370296; Pubmed=7642639;  
 RA Zhu X., Chang K.-H., He D., Mancini M.A., Brinkley W.R., Lee W.-H.;  
 RT "The C terminus of mitosis is essential for its nuclear localization, centromere/kinetochore targeting, and dimerization.";  
 RL J. Biol. Chem. 270:19545-19550(1995).  
 RN [5]  
 RP CHARACTERIZATION.  
 RA MEDLINE=98437347; Pubmed=9763420;  
 RA Chan G.K.T., Schaar B.T., Yen T.J.;  
 RT "Characterization of the kinetochore binding domain of CENP-E reveals interactions with the kinetochore proteins CENP-F and hBUBR1.";  
 RL J. Cell Biol. 143:49-63(1998).  
 CC -1- FUNCTION: PROBABLY REQUIRED FOR KINETOCHORE FUNCTION, INVOLVED IN CHROMOSOME SEGREGATION DURING MITOSIS. INTERACTS WITH RETINOBLASTOMA PROTEIN (RB), CENP-E AND BUBR1.  
 CC -1- SUBUNIT: HOMO- OR HETERODIMER.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX (BUT NOT IN THE NUCLEOLUS), REORGANIZATION TO THE KINETOCHORE/CENTROMERE (CORONAL SURFACE OF THE OUTER PLATE) AND THE SPINDLE DURING MITOSIS.  
 CC -1- DEVELOPMENTAL STAGE: GRADUALLY ACCUMULATES DURING THE CELL CYCLE.  
 CC -1- PTM: HYPERPHOSPHORYLATED DURING MITOSIS.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U19769; AAA82889.1; -;  
 DR EMBL; U30872; AAA82935.1; -;  
 DR EMBL; U25725; AAA86889.1; -;  
 DR HSSP; P02649; 1LE4.  
 DR MTM; 600236; -;  
 KW Chromosomal protein; Nuclear protein; Centromere; Coiled coil;  
 KM Mitosis; Phosphorylation; Antigen; Cell cycle; Repeat.  
 FT DOMAIN 14 197  
 FT DOMAIN 273 769 COILED COIL (POTENTIAL).  
 FT DOMAIN 823 1328 COILED COIL (POTENTIAL).  
 FT DOMAIN 1642 1746 COILED COIL (POTENTIAL).  
 FT DOMAIN 1862 2987 COILED COIL (POTENTIAL).  
 FT DOMAIN 2207 2568 COILED COIL (POTENTIAL).  
 FT REPEAT 2207 2386 2 X 177 AA TANDEM REPEATS.  
 FT REPEAT 2389 2568 1.  
 FT DOMAIN 3015 3032  
 FT DOMAIN 16 16 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT CONFLICT 250 250 L -> A (IN REF. 2).  
 FT CONFLICT 272 272 L -> Q (IN REF. 2).  
 FT CONFLICT 611 611 G -> D (IN REF. 2).  
 FT CONFLICT 611 611 MISSING (IN REF. 2).  
 FT CONFLICT 1494 1589 MISSING (IN REF. 2).  
 FT CONFLICT 1611 1611 V -> A (IN REF. 2).  
 FT CONFLICT 1811 1811 V -> L (IN REF. 2).  
 FT CONFLICT 2243 2243 ER -> DG (IN REF. 3).  
 FT CONFLICT 2335 2335 L -> Q (IN REF. 3).  
 FT CONFLICT 2492 2492 D -> N (IN REF. 2).  
 FT CONFLICT 2545 2561 ELNRYVALHNDQACK -> SSMREMPQIMTKKPYVS (IN REF. 3).  
 SQ SEQUENCE 3210 AA; 367589 MW; 11D83524960E4334 CRC64;

Query Match 14.5%; Score 155; DB 1; Length 3210;  
 Best Local Similarity 23.6%; Pred. No. 0.31; Mismatches 90; Indels 40; Gaps 7;  
 Matches 57; Conservative 55;  
 QY 19 LDREKNELDNVRALQSLQKREKRDQVITLRLDTLEERNATVVSLOALGKAEMLCSTLRKQ 77  
 DB 1204 LAEQIDMLQKVKQKLEKESEFKLEIDLSISVSE-----SVSKSKANLEKICRTLEDQ 1256

Db 2299 LDVLTARSEKENTKQIOEKOGOLSELDKLLSFKSLKEKQAEIQIKESKTAVENTMQ 2358  
OY 78 STLKKO-----MKYLPQOOD---ETKQAOEAGRLRSKKMTTEQTELLIQS 120  
Db 2359 NQKLENAVALGDDGEIMKATEOSLDPPLEEHQOLNTEKRLARLEADEKQOLCYLQ 2418  
OY 121 QLPPE-----VEEMIRDMVGOSAVBQALVAVCYSLKREYENLKARKASGEVADK 169  
Db 2419 QKRESEHHADLLKGRVENLEHLEIARTNOHALLEAVNSGVEVETLKAKTEGWTQSLRG 2478  
OY 170 LRKLFSSRSKLTQVYSELDQ-----AKLEL--KSQKDLQADKETSLLKK-----LTM 218  
Db 2479 LEIDVYTRSEKEDLTNELQKOEERISELEIINSSFENILQEKQKQKVKEMKESSTAMEM 2538  
OY 219 LQ 220  
Db 2539 LQ 2540

RESULT 16  
VDP\_RAT ID VDP\_RAT STANDARD: PRT: 959 AA.  
AC P1542:  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE General vesicular transport factor p15 (Transcytosis associated protein) (TAP) (Vesicle docking protein).  
CN VDP.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_Taxid=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=95132632; PubMed=7831323;  
RA Sapperstein S.K., Walter D.M., Grosvenor A.R., Heuser J.E., Waters M.G.;  
RT "p15 is a general vesicular transport factor related to the yeast endoplasmic reticulum to Golgi transport factor Uosp1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 92:522-526(1995).  
RN [2]  
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
RC TISSUE=Liver;  
RX MEDLINE=95132633; PubMed=7831324;  
RA Barroso M., Nelson D.S., Szul E.;  
RT "transcytosis-associated protein (TAP)/p15 is a general fusion factor required for binding of vesicles to acceptor membranes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 92:527-531(1995).  
CC -1- FUNCTION: GENERAL VESICULAR TRANSPORT FACTOR REQUIRED FOR INTERCISTERNAL TRANSPORT IN THE GOLGI STACK; IT IS REQUIRED FOR TRANSCYTOTIC FUSION AND/OR SUBSEQUENT BINDING OF THE VESICLES TO THE TARGET MEMBRANE. MAY WELL ACT AS A VESICULAR ANCHOR BY INTERACTING WITH THE TARGET MEMBRANE AND HOLDING THE VESICULAR AND TARGET MEMBRANES IN PROXIMITY.  
CC -1- SUBUNIT: HOMODIMER WITH TWO N-TERMINAL HEADS AND A C-TERMINAL COILED-COIL TAIL. DIMER FORMED BY PARALLEL ASSOCIATION OF THE TAILS.  
CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN WHICH RECYCLES BETWEEN THE CYTOSOL AND THE GOLGI APPARATUS DURING INTERPHASE.  
CC -1- DOMAIN: COMPOSED OF A GLOBULAR HEAD, AN ELONGATED TAIL (COILED-COIL) AND A HIGHLY ACIDIC C-TERMINAL DOMAIN.  
CC -1- PTM: PHOSPHORYLATED IN A CELL CYCLE-SPECIFIC MANNER.  
CC PHOSPHORYLATED IN INTERPHASE BUT NOT IN MITOTIC CELLS.  
CC DEPHOSPHORYLATED PROTEIN ASSOCIATES WITH THE GOLGI MEMBRANE;  
CC PHOSPHORYLATED PROMOTES DISSOCIATION (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE VDP/OSOI/YELO4 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U14192; AA62632.1; -;  
DR EMBL: U15589; AAC52151.1; -;  
DR InterPro: IPR000225; Armadillo.  
DR PROSITE: PS50176; ARM\_REPEAT; UNKNOWN\_1.  
KW Transport; Protein transport; Golgi stack; Membrane; Coiled coil;  
KW Phosphorylation.  
FT DOMAIN 1 637 GLOBULAR HEAD.  
FT DOMAIN 638 930 COILED COIL (POTENTIAL).  
FT DOMAIN 935 959 ASP/GLU-RICH (ACIDIC).  
FT MOD\_RES 940 940 PHOSPHORYLATION (BY SIMILARITY).  
FT CONFLICT 591 591 S -> P (IN REF. 2).  
FT CONFLICT 658 658 M -> V (IN REF. 2).  
FT CONFLICT 816 816 S -> R (IN REF. 2).  
FT CONFLICT 873 873 A -> S (IN REF. 2).  
SQ SEQUENCE 959 AA; 107162 MW; 356394B48C7E003B CRC64;

Query Match 14.5%; Score 154.5; DB 1; Length 959;  
Best Local Similarity 20.1%; Pred. No. 0.097;  
Matches 66; Conservative 53; Mismatches 78; Indels 131; Gaps 9;  
OY 15 EENVLDREFLK--NELDNV--RAQLSQKDKKRSQVITPLRDTLEERNATVVSLOAL 70  
Db 605 EYMFDFHEFTKLVKELEGVITKATIKKSEDEKKEEV-----KTLQHDNIYVHYKNMI 659  
OY 71 GKAEMLCSTLKKQKKYLEQODE-----TKQAO----- 98  
Db 660 REQLQLEELKQVSTLKQNEQVLTQAVTQOASQIQGHKQYNLKVQLGKDNHIGSHS 719  
OY 99 -----FEAGRLSKSKKMTQEIIELLOSLQSPVPEBMRDM-----GVGSAV--- 139  
Db 720 DGAQVNGIQPEISRLREELRSHQVLLQSLAEKDVYENLRSSVGSMSBALATC 779  
OY 140 -----EQLAVY----- 145  
Db 780 SPRADQVAVELKQELSLKSLQSLCSQSLIEITRLQENSELDQARATLAKSVPEGESELYT 839  
OY 146 -----CVSKREYENLKARKASGEVADKLRDLFSSRSKLTQVYSELDQKLEL 195  
Db 840 AAKTVDVEGRSLAQETKELKNEIKALSEREATIQOLDSSNSTIATLQTEKDKLYLEV 899  
OY 196 KSQKQD-----LQSADEKETSLLKK 216  
Db 900 TDSKKEQDDLLVLTADQDKILSKSL 927

RESULT 17  
MYSN\_ACACA ID MYSN\_ACACA STANDARD: PRT: 1509 AA.  
AC P05659;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Myosin II heavy chain, non muscle.  
OS Acanthamoeba castellanii (Amoeba).  
OC Eukaryota; Acanthamoebidae; Acanthamoeba.  
OX NCBI\_Taxid=5755;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87308395; PubMed=3040773;  
RA Hammer J.A. III, Bowers B., Paterson B.M., Korn E.D.;  
RT "Complete nucleotide sequence and deduced polypeptide sequence of a rat nonmuscle myosin heavy chain gene from Acanthamoeba: evidence of a hinge in the rodlike tail.";  
RL J. Cell Biol. 105:913-925(1987).  
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE ACTIVITY THAT IS ACTIVATED BY F-ACTIN.  
CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI

```

CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC (MLC-2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS
CC INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE
CC THE REGULATORY PHOSPHORYLATION SITES RESIDE.
CC -1- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING
CC THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMATIC ACTIVITY.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 IO DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y00624; CAA68663.1; -.
DR PIR: A27224; A27224.
DR HSSP: P08799; 1MND.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR004009; myosin_N.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 2.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; myosin_N; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
DR Myosin; Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;
KW Methylation; Alkylation; Phosphorylation; Multigene family.
FT DOMAIN 1 789
FT DOMAIN 790 819
FT DOMAIN 848 1509
FT DOMAIN 848 1226
FT DOMAIN 1227 1252
FT DOMAIN 1253 1509
FT DOMAIN 1253 1482
FT DOMAIN 1483 1509
FT NP_BIND 182 189
FT DOMAIN 660 682
FT DOMAIN 766 780
FT MOD_RES 133 133
FT MOD_RES 700 700
FT MOD_RES 1489 1489
FT MOD_RES 1494 1494
FT MOD_RES 1499 1499
SQ SEQUENCE 1509 AA; 171201 MW; 2CE49BE51173D17E CRC64;

Query Match 14.5%; Score 154.5; DB 1; Length 1509;
Best Local Similarity 25.8%; Pred. No. 0.15;
Matches 63; Conservative 52; Mismatches 91; Indels 55; Gaps 7;

```

```

DB 1074 QLEBEE---KSGREAASSKAKOLGOLEDARSEVLSKSLSAEKSLKTKAKDQNRDLDE 1129
QY 204 -----SADKEIMSLKKILTMLO 220
DB 1130 QLEDETVRANVDKOKALEKLELE 1156

RESULT 18
MYH2_HUMAN STANDARD; PRT; 1941 AA.
ID Q9UKX2; Q16229; Q14322;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, skeletal muscle, adult 2 (Myosin heavy chain I1a)
DE (MYHC-I1a).
GN MYH2 OR MYHSA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=9931869; PubMed=10388558;
RA Weiss A., Schiaffino S., Leinwand L.A.;
RT "Comparative sequence analysis of the complete human sarcomeric myosin
RT heavy chain family: implications for functional diversity.";
RL J. Mol. Biol. 290:61-75(1999).
RN [2]
RP SEQUENCE OF 1711-1941 FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=95109625; PubMed=7545970;
RA Smerud V., Karsch-Mizrachi I., Campione M., Leinwand L.,
RA Schiaffino S.;
RT "Type IIX myosin heavy chain transcripts are expressed in type IIB
RT fibers of human skeletal muscle.";
RL Am. J. Physiol. 267:C1723-C1728(1994).
RN [3]
RP SEQUENCE OF 1823-1941 FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=95270723; PubMed=7751403;
RA Emilion S., Sant'ana Pereira J., Sargeant T., Young A., Goldspink G.;
RT "Characterization of human skeletal muscle fibres according to the
RT myosin heavy chains they express.";
RL J. Muscle Res. Cell Motil. 16:35-43(1995).
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROTOSIN (LMH) AND 1 HEAVY MEROTOSIN (HMH). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF111784; AAD29950.1; -.
DR EMBL: S73840; AAC13916.1; -.

```

DR EMBL; 232858; CAA83687.1; -  
 DR HSSP; P08799; IMND.  
 DR MIM; 160740; -  
 DR InterPro; IPR000048; IQ.  
 DR InterPro; IPR004009; Myosin\_N.  
 DR InterPro; IPR002928; Myosin\_tail.  
 DR InterPro; IPR001609; myosin\_head.  
 DR Pfam; PF00612; IQ; 2.  
 DR Pfam; PF00063; myosin\_head; 1.  
 DR Pfam; PF02736; Myosin\_N; 1.  
 DR Pfam; PF01576; Myosin\_tail; 1.  
 DR PRINTS; PR00193; MYOSINHEAVY.  
 DR PRODOM; PD000355; myosin\_head; 1.  
 DR SMART; SM00015; IQ; 1.  
 DR SMART; SM00242; MYSC; 1.  
 DR PROSITE; PSS0096; IQ; 1.  
 DR Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;  
 KM Multigene family.  
 FT DOMAIN 1 786 MYOSIN HEAD-LIKE.  
 FT DOMAIN 787 816 IQ.  
 FT DOMAIN 845 1941 COILED COIL (POTENTIAL).  
 FT NP\_BIND 179 186 ATP (POTENTIAL).  
 FT CONFLICT 1844 1844 K -> R (IN REF. 3).  
 FT SEQUENCE 1941 AA; 223043 MW; 681E866F83AE83F CRC64;

Query Match 14.4%; Score 154; DB 1; Length 1941;  
 Best Local Similarity 23.3%; Pred. No. 0.21;  
 Matches 53; Conservative 51; Mismatches 79; Indels 44; Gaps 7;

QY 24 LKNEIDNRAQLSQKDKRDSQVIIDTLRLTEERNATVSLQALGKAEMLSTLTKQ 83  
 Db 1209 LGEQIDNLRKQKLEKSEKMEIDLASNV-----TSKAKNLEKCTLEDQ 1261  
 QY 84 MKYLEQODETK-----QAQEEAGRL-----RSKAKTMEQIELL 117  
 Db 1262 LSELSKSEKEEQRLINDLTAORGLQTESGEFSROLDEKALVSLRGKQAFQOIEL 1321  
 QY 118 LOSQLPEVEEMIRDMGVGQSAVEQLAVYCVSLKKEYLEMKARRASGEVAKLKRDLESS 177  
 Db 1322 KR----QLEEEIKAKNALAHLSQSRHDCDLIREQYEEQSKALELQALSKANTEVAQW 1377  
 QY 178 RSKLQT---YSELDAQLELKSQKDLQSADEKIMSLKKKLTMLQ 220  
 Db 1378 RTKYTEIDAIQRTLELEAKKTL-AQR-LQAAEHVEAVNAKCSLE 1421

RESULT 19  
 G160\_HUMAN STANDARD; PRT; 579 AA.  
 AC 008378;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Golgin-160 (Fragment).  
 GN GOLGA3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=93301617; PubMed=8315394;  
 RA Fitzler M.J., Hamel J.C., Ochs R.L., Chan E.K.L.;  
 RT "Molecular characterization of two human autoantigens: unique CDNA  
 RT encoding 95- and 160-kD proteins of a putative family in the golgi  
 RT complex";  
 RL J. Exp. Med. 178:49-62(1993).  
 CC -1- FUNCTION: GOLGI AUTO-ANTIGEN. MAY HAVE A FUNCTION IN THE  
 CC PROCESSING AND TRANSPORT OF PROTEINS THROUGH THE GOLGI.  
 CC -1- SIMILARITY: HIGH, TO MOUSE MALE-ENHANCED ANTIGEN-2 (MEA-2).

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L06148; AAA35921.1; -  
 DR HSSP; P80220; IDIP.  
 DR MIM; 602581; -  
 DR Golgi stack.  
 FT NON\_TER 1 1  
 FT NON\_TER 579 579  
 FT SEQUENCE 579 AA; 66350 MW; 8C9A5A73D0128AEE CRC64;

Query Match 14.3%; Score 152.5; DB 1; Length 579;  
 Best Local Similarity 25.0%; Pred. No. 0.075;  
 Matches 55; Conservative 42; Mismatches 86; Indels 37; Gaps 6;

QY 13 QEEENVLDREFLKNELD-----NVRQLSQKDEKRDQVIIDTLRLTEERNATVSL 66  
 Db 363 REHNSILETALAKREADVQLNLQVAVLQVRKEEDEROMKHLVQALQASLEKEKYNVL 422  
 QY 67 QQALGKAEMLCSTLKKKKYLEQODETK-----QAQEEAGRSLSKMTME 112  
 Db 423 KEQYAAKAVENGHNRRHKAASLSEYKKEQLQAKKHLVQKQLQAEADLQIREGKHSQEI 482  
 QY 113 ---QIEL-----LQSQLPVEVEEMIRDMGVGQSAVEQLAVYCVSLKKEYLEMKARK 161  
 Db 483 AQQAELAEARAOQLQQLQKQ-----DEQLSKQPVGNQENKMEVDQKEREIQLKQQLD 539  
 QY 162 ASGEVADKLKRDLFSSRSKLTQTVYSELDAQLELKSQKDL 201  
 Db 540 LTEQOG---RKLEGLQDLQNLQNVKSELEMAQEDLSMTQKD 576

RESULT 20  
 MYH7\_HUMAN STANDARD; PRT; 1935 AA.  
 AC P12883; Q14904; Q16579;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Myosin heavy chain, cardiac muscle beta isoform (MyHC-beta).  
 GN MYH7 OR MYHCB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90301496; PubMed=2249844;  
 RA Jaenike T., Diederich K.W., Haas W., Schleich J., Lichter P.,  
 RA Pfordt M., Bach A., Vosberg H.P.;  
 RT "The complete sequence of the human beta-myosin heavy chain gene and  
 RT a comparative analysis of its product";  
 RL Genomics 8:194-206(1990).  
 RN 12  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90301496; PubMed=2362820;  
 RA Liew C.-C., Sole M.J., Yamauchi-Takahara K., Kellam B.,  
 RA Anderson D.H., Lin L., Liew J.;  
 RT "Complete sequence and organization of the human cardiac beta-myosin  
 RT heavy chain gene";  
 RL Nucleic Acids Res. 18:3647-3651(1990).  
 RN 13  
 RP SEQUENCE OF 1-115 FROM N.A.  
 RX MEDLINE=89264452; PubMed=2726733;  
 RA Yamauchi-Takahara K., Sole M.J., Liew J., Ing D., Liew C.-C.;  
 RT "Characterization of human cardiac myosin heavy chain genes";

RL Proc. Natl. Acad. Sci. U.S.A. 86:3504-3508(1989).  
 RN [14]  
 RN ERRATUM.  
 RA Yamauchi-Takahara K., Sole M.J., Liew J., Ing D., Liew C.-C.;  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:7416-7417(1989).  
 RN [15]  
 RN SEQUENCE OF 1310-1935 FROM N.A.  
 RX MEDLINE=86176778; PubMed=2421254;  
 RA Saez L., Leinwand L.A.;  
 RT "Characterization of diverse forms of myosin heavy chain expressed in  
 RL adult human skeletal muscle.";  
 RN Nucleic Acids Res. 14:2951-2959(1986).  
 RP REVISIONS.  
 RA Leinwand L.A.;  
 RL Submitted (MAR-1988) to the EMBL/GenBank/DBJ databases.  
 RN [17]  
 RN SEQUENCE OF 1410-1935 FROM N.A.  
 RX MEDLINE=88299163; PubMed=2969919;  
 RA Kurabayashi M., Tsuchimochi H., Komuro I., Takaku F., Yazaki Y.;  
 RT "Molecular cloning and characterization of human cardiac alpha- and  
 RT beta-form myosin heavy chain complementary DNA clones. Regulation of  
 RT expression during development and pressure overload in human  
 RT atrium.";  
 RL J. Clin. Invest. 82:524-531(1988).  
 RN [8]  
 RN SEQUENCE OF 785-1935 FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=90235862; PubMed=1691980;  
 RA Bobber E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,  
 RA Arnold H.H.;  
 RT "Identification of three developmentally controlled isoforms of human  
 RT myosin heavy chains.";  
 RL Eur. J. Biochem. 189:55-65(1990).  
 RN [9]  
 RN SEQUENCE OF 1393-1935 FROM N.A.  
 RX MEDLINE=87192738; PubMed=3032769;  
 RA Jandreski M.A., Liew C.-C.;  
 RT "Construction of a human ventricular cDNA library and  
 RT characterization of a beta myosin heavy chain cDNA clone.";  
 RL Hum. Genet. 76:47-53(1987).  
 RN [10]  
 RN REVIEW ON VARIANTS.  
 RP MEDLINE=96039076; PubMed=8533830;  
 RA Arai S., Matsuo K., Hirayama K., Sukurai H., Tamura M., Ozawa T.,  
 RA Kimura M., Imanura S.-I., Furutani Y., Joh-O K., Kawana M., Takao A.,  
 RA Hosoda S., Momma K.;  
 RT "Missense mutation of the beta-cardiac myosin heavy-chain gene in  
 RT hypertrophic cardiomyopathy.";  
 RL Am. J. Med. Genet. 58:267-276(1995).  
 RN [11]  
 RN VARIANTS CMH1 GLN-256 AND ARG-741.  
 RX MEDLINE=93248216; PubMed=8483915;  
 RA Pananapazir L., Dalakas M.C., Cyran F., Cohn G., Epstein N.D.;  
 RT "Missense mutations in the beta-myosin heavy-chain gene cause central  
 RT core disease in hypertrophic cardiomyopathy.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:3993-3997(1993).  
 RN [12]  
 RN VARIANT CMH1 GLN-403.  
 RX MEDLINE=90367131; PubMed=1975517;  
 RA Geisler-Lowrance A.A.T., Kass S., Tanigawa G., Vosberg H.-P.,  
 RA McKenna W., Seidman C.E., Seidman J.G.;  
 RT "A molecular basis for familial hypertrophic cardiomyopathy: a beta  
 RT cardiac myosin heavy chain gene missense mutation.";  
 RL Cell 62:999-1006(1990).  
 RN [13]  
 RN VARIANTS CMH1 GLN-249; GLN-403; CYS-453 AND MET-606.  
 RX MEDLINE=92204193; PubMed=1552912;  
 RA Watkins H., Rosenzweig A., Hwang D.S., Levi T., McKenna W.,  
 RA Seidman C.E., Seidman J.G.;  
 RT "Characteristics and prognostic implications of myosin missense  
 RT mutations in familial hypertrophic cardiomyopathy.";  
 RL New Engl. J. Med. 326:1108-1114(1992).

RN [14]  
 RN VARIANTS CMH1 GLN-403; CYS-453; ARG-584 AND MET-606.  
 RX MEDLINE=94070863; PubMed=8250038;  
 RA Watkins H., Thierfelder L., Anan R., Jarcho J., Matsumori A.,  
 RA McKenna W., Seidman J.G., Seidman C.E.;  
 RT "Independent origin of identical beta cardiac myosin heavy-chain  
 RT mutations in hypertrophic cardiomyopathy.";  
 RL Am. J. Hum. Genet. 53:1180-1185(1993).  
 RN [15]  
 RN VARIANTS CMH1 GLN-403 AND VAL-908.  
 RX MEDLINE=92346810; PubMed=1638703;  
 RA Epstein N.D., Cohn G.M., Cyran F., Pananapazir L.;  
 RT "Differences in clinical expression of hypertrophic cardiomyopathy  
 RT associated with two distinct mutations in the beta-myosin heavy chain  
 RT gene. A 908Leu->Val mutation and a 403Arg->Gln mutation.";  
 RL Circulation 86:345-352(1992).  
 RN [16]  
 RN VARIANTS CMH1 LEU-403 AND TRP-403.  
 RX MEDLINE=94075629; PubMed=8254035;  
 RA Dausse E., Komajda M., Feller L., Dubourg O., Dufour C., Carrier L.,  
 RA Wansky C., Bericovic J., Hengstenberg C., Al-Mahdawi S.;  
 RT "Familial hypertrophic cardiomyopathy. Microsatellite haplotyping and  
 RT identification of a hot spot for mutations in the beta-myosin heavy  
 RT chain gene.";  
 RL J. Clin. Invest. 92:2807-2813(1993).  
 RN [17]  
 RN VARIANT CMH1 TRP-403.  
 RX MEDLINE=94093568; PubMed=8268932;  
 RA Woolman J.C., Brink P.A., Corfield V.A.;  
 RT "Identification of a new missense mutation at Arg403, a Cpg mutation  
 RT hotspot, in exon 13 of the beta-myosin heavy chain gene in  
 RT hypertrophic cardiomyopathy.";  
 RL Hum. Mol. Genet. 2:1731-1732(1993).  
 RN [18]  
 RN VARIANT CMH1 ASN-615.  
 RX MEDLINE=93038688; PubMed=1417858;  
 RA Nishi H., Kimura A., Harada H., Yoshima H., Sasazuki T.;  
 RT "Novel missense mutation in cardiac beta myosin heavy chain gene  
 RT found in a Japanese patient with hypertrophic cardiomyopathy.";  
 RL Biochem. Biophys. Res. Commun. 188:379-387(1992).  
 RN [19]  
 RN VARIANT CMH1 GLY-778.  
 RX MEDLINE=93343938; PubMed=8343162;  
 RA Harada H., Kimura A., Nishi H., Sasazuki T., Yoshima H.;  
 RT "A missense mutation of cardiac beta-myosin heavy chain gene linked  
 RT to familial hypertrophic cardiomyopathy in affected Japanese  
 RT families.";  
 RL Biochem. Biophys. Res. Commun. 194:791-798(1993).  
 RN [20]  
 RN VARIANT CMH1 VAL-908.  
 RX MEDLINE=93168485; PubMed=8435239;  
 RA Al-Mahdawi S., Chamberlain S., Cleland J., Ninoyannopoulos P.,  
 RA Gilligan D., French J., Choudhury L., Williamson R., Oakley C.;  
 RT "Identification of a mutation in the beta cardiac myosin heavy chain  
 RT gene in a family with hypertrophic cardiomyopathy.";  
 RL Br. Heart J. 69:136-141(1993).  
 RN [21]  
 RN VARIANT CMH1 TRP-719.  
 RX MEDLINE=95179132; PubMed=7874131;  
 RA Greve G., Bachinski L., Friedman D.L., Czernezewicz G., Anan R.,  
 RA Towbin J., Seidman C.E., Roberts R.;  
 RT "Isolation of a de novo mutant myocardial beta MHC protein in a  
 RT pedigree with hypertrophic cardiomyopathy.";  
 RL Hum. Mol. Genet. 3:2073-2075(1994).  
 RN [22]  
 RN VARIANTS CMH1 CYS-513; ARG-716 AND TRP-719.  
 RX MEDLINE=94110336; PubMed=8282798;  
 RA Anan R., Greve G., Thierfelder L., Watkins H., McKenna W., Solomon S.,  
 RA Vecchio C., Shono H., Nakao S., Tanaka H., Mares A. Jr., Towbin J.A.,  
 RA Spirito P., Roberts R., Seidman J.G., Seidman C.E.;  
 RT "Prognostic implications of novel beta cardiac myosin heavy chain gene  
 RT mutations that cause familial hypertrophic cardiomyopathy.";  
 RL J. Clin. Invest. 93:280-285(1994).



RN [23]  
 RP VARIANT CMH1 THR-797.  
 RX MEDLINE=96047159; PubMed=7581410;  
 RA Moolman J.C., Brink P.A., Corfield V.A.;  
 RT "Identification of a novel Ala797Thr mutation in exon 21 of the beta-  
 RL myosin heavy chain gene in hypertrophic cardiomyopathy.";  
 RN Hum. Mutat. 6:197-198(1995).  
 RN [24]  
 RP VARIANT CMH1 CYS-453.  
 RX MEDLINE=96209901; PubMed=8655135;  
 RA Ko Y.-L., Chen J.-J., Tang T.-K., Cheng J.-J., Lin S.-Y., Liou Y.-C.,  
 RA Kuan P., Wu C.-W., Lien W.-P., Liew C.-C.;  
 RT "Major familial hypertrophic cardiomyopathy in a family with a  
 RT 453Arg->Cys mutation in the beta-myosin heavy chain gene;  
 RT coexistence of sudden death and end-stage heart failure.";  
 RL Hum. Genet. 97:585-590(1996).  
 RN [25]  
 RP VARIANTS CMH1 THR-349 AND TRP-719.  
 RX MEDLINE=98204402; PubMed=9544842;  
 RA Jesche B., Uhl K., Weist B., Schröder D., Meitinger T.,  
 RA Dohlemann C., Vosberg H.-P.;  
 RT "A high risk phenotype of hypertrophic cardiomyopathy associated with  
 RT Query Match 14.3%; Score 152.5; DB 1; Length 1935;  
 Best Local Similarity 20.5%; Pred. No. 0.25;  
 Matches 61; Conservative 60; Mismatches 88; Indels 89; Gaps 9;  
 QY 1 RTIINKLEFDLAOEENLDR-----EFLK----- 25  
 DB 1129 RAKVEKLSDSLRELEISELEEGATSVQIEMKKREAFQMRDLEATLQHEAT 1188  
 QY 26 -----NELDNVRAQLSOKDEKRDQVITITLRTLEERNATVYLSQA 69  
 DB 1189 AALRLKKHADSVAELGIEDNLQRYKQKLEKESEFKLEDDVSNMQ-----IKKA 1241  
 QY 70 LGAEMLCSTLKKQMKYLEDQOODETKQAQEAERLSRMKT-----MEQIELLSQ 121  
 DB 1242 KANLEKMCRTLEDQNEHRSKAEETORSVNDLTSQRAKLQENGLSHQDEKELI-SQ 1300  
 QY 122 LP-----EVEEMIRMGVGSADVEDLAVYCVSLKKEVENLKEARASEV 166  
 DB 1301 LTRGKLYTQOLEDLKRQLEBEVAKNALAHALQARHDCDLRQRYEETLEAKAEIQRV 1360  
 QY 167 ADLRLKDLFSSRSKLTQ-----VYSELDAQLELKSQKDLQADKEIMSLKKITMLQ 220  
 DB 1361 LSKANSEVAQMRKYETDAIQRTLELEAKKL--AGR-LQEAEEAVANAKCSLE 1415  
 RESULT 21  
 MYSA\_DROME STANDARD: PRT; 2411 AA.  
 AC POS661:  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Myosin heavy chain, muscle.  
 GN MHC.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89384536; PubMed=2506434;  
 RA George E.L., Ober M.B., Emerson C.P. Jr.;  
 RT "Functional domains of the Drosophila melanogaster muscle myosin  
 RT heavy-chain gene are encoded by alternatively spliced exons.";  
 RL Mol. Cell. Biol. 9:2957-2974(1989).  
 RN [2]  
 RP SEQUENCE OF 1-312 FROM N.A.  
 RX MEDLINE=87280141; PubMed=3038896;  
 RA Wassenberg D.R. II, Kronert W.A., O'Donnell P.T., Bernstein S.I.;

RT "Analysis of the 5' end of the Drosophila muscle myosin heavy chain  
 RT gene. Alternatively spliced transcripts initiate at a single site and  
 RT intron locations are conserved compared to myosin genes of other  
 RT organisms.";  
 RL J. Biol. Chem. 262:10741-10747(1987).  
 RN [3]  
 RP SEQUENCE OF 486-881 FROM N.A.  
 RC STRAIN-CANTON-S; TISSUE=Embryonic muscle;  
 RX MEDLINE=91330870; PubMed=1907912;  
 RA Kronert W.A., Edwards K.A., Roche E.S., Wells L., Bernstein S.I.;  
 RT "Muscle-specific accumulation of Drosophila myosin heavy chains: a  
 RT splicing mutation in an alternative exon results in an isoform  
 RT substitution.";  
 RL EMBO J. 10:2479-2488(1991).  
 CC -1- FUNCTION: MUSCLE CONTRACTION.  
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -1- ALTERNATIVE PRODUCTS: AS DROSOPHILA HAS A SINGLE MUSCLE MHC GENE,  
 CC MHC ISOFORMS MUST RELY ON ALTERNATIVE RNA SPLICING THAT LEADS TO  
 CC DIFFERENCES IN THE C-TERMINUS OF THE VARIOUS MHC PROTEINS.  
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M61229; AAA28686.1; ALT\_SEQ.  
 DR EMBL: M61229; AAA28687.1; ALT\_SEQ.  
 DR EMBL: J02788; AAA28706.1; ALT\_SEQ.  
 DR EMBL: J02788; AAA28707.1; ALT\_SEQ.  
 DR EMBL: X60196; CAA42752.1; ALT\_SEQ.  
 DR EMBL: X60196; CAA42753.1; ALT\_SEQ.  
 DR EMBL: X60196; CAA42754.1; ALT\_SEQ.  
 DR PIR: A28492; A28492.  
 DR PIR: A32491; A32491.  
 DR PIR: B32491; B32491.  
 DR HSP: P08799; LMND.  
 DR FLYBase: FBgn0002741; Mhc.  
 DR InterPro: IPR000048; IQ.  
 DR InterPro: IPR004009; Myosin\_N.  
 DR InterPro: IPR002928; Myosin\_tail.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00612; IQ; 2.  
 DR Pfam: PF00063; myosin\_head; 7.  
 DR Pfam: PF02736; Myosin\_N; 1.  
 DR Pfam: PF01576; Myosin\_tail; 1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR PRODom: PD000335; myosin\_head; 5.  
 DR SMART: SM00015; IQ; 1.  
 DR SMART: SM00242; MYSC; 1.  
 DR PROSITE: PS50096; IQ; 1.  
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KW ATP-binding; Calmodulin-binding; Alternative splicing;  
 KW Multigene family.  
 FT DOMAIN 1 ? MYOSIN HEAD-LIKE.  
 FT DOMAIN 1205 1232 IO.  
 FT DOMAIN 1233 2376 COILED COIL (POTENTIAL).  
 FT NP\_BIND 227 234 ATP (BY SIMILARITY).  
 FT VARSPPLIC 2385 2385 P->I (IN SHORT ISOFORM).  
 FT VARSPPLIC 2386 2411 MISSING (IN SHORT ISOFORM).  
 FT CONFLICT 43 44 EK -> RE (IN REF. 2).  
 SQ SEQUENCE 2411 AA; 276434 MW; 8C4FEFC1BB8D310A9 CRC64;



DR InterPro: IPR002928; Myosin\_tail.  
DR InterPro: IPR001609; myosin\_head.  
DR Pfam: PF00612; IQ: 1.  
DR Pfam: PF00063; myosin\_head; 1.  
DR Pfam: PF02736; myosin\_N; 1.  
DR Pfam: PF01576; Myosin\_tail; 1.  
DR PRINTS: PR00193; MYOSINHEAVY.  
DR PRODOM: PD000355; myosin\_head; 1.  
DR SMART: SM00015; IQ: 1.  
DR SMART: SM00242; MYSC; 1.  
DR PROSITE: PS50096; IQ: 1.  
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
KW ATP-binding; Methylation; Alkylation; Multigene family.  
FT DOMAIN 1 793  
FT DOMAIN 794 823  
FT NP\_BIND 179 186 COILED COIL (POTENTIAL).  
FT DOMAIN 667 689 ATP-BINDING.  
FT DOMAIN 770 784 ACTIN-BINDING.  
FT MOD\_RES 130 130 METHYLATION (TRI-) (POTENTIAL).  
FT MOD\_RES 707 707 ALKYLATION (SH-1) (POTENTIAL).  
FT MOD\_RES 717 717 ALKYLATION (SH-2) (POTENTIAL).  
SQ SEQUENCE 1969 AA: 225509 MW: 64577BBAF7EAD80A CRC64;

Query Match 14.1%; Score 150; DB 1; Length 1969;  
Best Local Similarity 23.5%; Pred. No. 0.35;  
Matches 64; Conservative 48; Mismatches 94; Indels 66; Gaps 10;

OY 5 NKLFPDLAQBENVLDRFLKNELDNVAQLSQK-----DKEKRD 45  
DB 898 NALFSLSETEKANLADAEERNEKLNQLKATLESKLSIDITGQLEDWQERNEDLARKKKTD 957  
OY 46 QVIIDT-----LBDTLEE---RNATVVSQQALGKAKMCSITLKKQKMTLEOOD 92  
DB 958 QELSTQKKHVDLELSLTKKABQKSDRNIRSLQDENANQDEAVAKLNKREKH---QEE 1014  
OY 93 ETQKQAEAGRLSKMTMEQIELLQSLQPEVEEMI---RDMGVGSAVEQLAVYCVS 148  
DB 1015 SNRKINDELQSEEDKVNLEKIRNKLEQMDLEENIDREKSRSDIKARKRYE---GD 1071  
OY 149 LKKEYEINKEARKASGEVADKLK---DLFSSR-----SKLQTVYSEL----- 188  
DB 1072 LKVAQENIDETLKQKHVDETLLKREEDLHNTNAKLAENNSIAKQLQILKELFARNAL 1131  
OY 189 -DQAKLELKSNOXQDS---ADKEIMSLKKTL 216  
DB 1132 EEELEAERNRQKSDRSRSEAELELELTERL 1163  
RESULT 24  
KUCR\_MOUSE STANDARD: PRT: 548 AA.  
ID KUCR\_MOUSE STANDARD: PRT: 548 AA.  
AC P70194;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Kuppfer cell receptor.  
GN KCLR.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C; TISSUE=Liver;  
RA Takekawa R., Magatsuma H., Nomoto C., Watanabe Y., Akaike T.;  
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: RECEPTOR WITH AN AFFINITY FOR GALACTOSE AND FUCOSE.  
CC COULD BE INVOLVED IN ENDOCYTOSIS.  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.  
CC -1- TISSUE SPECIFICITY: KUPFER CELLS.  
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
CC -----

DR EMBL: D88577; BA113647.1; -.  
DR HSSP: P20693; HLJ.  
DR MGD: MGI:1859834; Kclr.  
DR InterPro: IPR001304; Lectin\_c.  
DR Pfam: PF00059; lectin\_c; 1.  
DR SMART: SM00034; LECTC; 1.  
DR PROSITE: PS00615; C-TYPE\_LECTIN\_1; 1.  
DR PROSITE: PS50041; C-TYPE\_LECTIN\_2; 1.  
KW Receptor; Transmembrane; Glycoprotein; Lectin; Signal-anchor;  
KW Endocytosis.  
FT DOMAIN 1 42 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 43 69 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT FT  
FT DOMAIN 438 548 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 440 536 C-TYPE LECTIN (SHORT FORM).  
FT DISULFID 516 528 BY SIMILARITY.  
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 548 AA: 61268 MW: 6F6495E820E73BD9 CRC64;

Query Match 14.0%; Score 149; DB 1; Length 548;  
Best Local Similarity 26.0%; Pred. No. 0.11;  
Matches 59; Conservative 45; Mismatches 85; Indels 38; Gaps 9;

OY 22 EFLKNELDNVAQLSQ-KDKEKRSQVITDR-LRDTLEERNATVVSQQALGKAKMCS 78  
DB 138 QLTGSHELDVANADILQTDVLKESGALALETQALRSSLEVASADHSIRGDLKANAMTS 197  
OY 79 TLKQMKY-----LEQOQDETQKQAEAGRLSKKMTMEQIELLQSLQ 123  
DB 198 QTRGLKSTENTSANELHVLGRGLEAQS---IQALRGSQASANDLSSQTGFLQHSMD 254  
OY 124 EEEEMIRDMGVGQSAVEQLAVYCVSLKREYENL-KEARKASGEVADK-----LRKDLFS 176  
DB 255 NISAQIQTV---RDMERAGKKNMSLKKELETLTAQOTKANGHLEQTDQIQGLKAEKLS 311  
OY 177 S---RSKIQTVYSELQDAKLELKSNOQDQASADEIKSLKKLIMQ 220  
DB 312 TSSUNSLRIEYVNGMKDASRELQTLRRDL---SDVSALKSNVOMLQ 354  
RESULT 25  
RAS0\_AERPE STANDARD: PRT: 919 AA.  
ID RAS0\_AERPE STANDARD: PRT: 919 AA.  
AC Q9YFZ1;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE DNA double-strand break repair rad50 ATPase.  
GN RAD50 OR APE0110.  
OS Aeropyrum pernix.  
OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;  
OX Aeropyrum.  
OX NCBI\_TaxID=56636;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KI;  
RA MEDLINE=99310339; PubMed=10382966;



RC TISSUE=Skeletal muscle;  
 RX MEDLINE=99318869; PubMed=10388558;  
 RA Weiss A., Schiaffino S., Leinwand L.A.;  
 RT "Comparative sequence analysis of the complete human sarcomeric myosin  
 heavy chain family: implications for functional diversity.";  
 RL J. Mol. Biol. 290:61-75(1999).  
 RN [2]  
 RP SEQUENCE OF 1064-1939 FROM N.A.  
 RA MEDLINE=8617678; PubMed=2421254;  
 RX Saez L., Leinwand L.A.;  
 RT "Characterization of diverse forms of myosin heavy chain expressed in  
 adult human skeletal muscle."  
 RL Nucleic Acids Res. 14:2951-2969(1986).  
 CC -1- FUNCTION: MUSCLE CONTRACTION.  
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -1- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1- PPM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 AKTYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE  
 SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 SUBFRAGMENT (S2).  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF111785; AAD29951.1; -  
 DR EMBL: X03740; CAA27380.1; -  
 DR PIR: A23767; A23767.  
 DR HSSP: P03437; 1HTM.  
 DR MIM: 160730; -  
 DR InterPro: IPR000048; IO.  
 DR InterPro: IPR002928; Myosin\_tail.  
 DR InterPro: IPR001609; Myosin\_head.  
 DR Pfam: PF00063; myosin\_head.1.  
 DR Pfam: PF01576; myosin\_tail.2.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR PRODOM: PD000355; myosin\_head.1.  
 DR SMART: SM00015; IO.1.  
 DR SMART: SM00242; MYSC.1.  
 DR PROSITE: PS50096; IO.1.  
 DR Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;  
 KW Multigene family.  
 FT DOMAIN 1 784 MYOSIN HEAD-LIKE.  
 FT DOMAIN 785 814 IO.  
 FT DOMAIN 843 1939 COILED COIL (POTENTIAL).  
 FT NP\_BIND 179 186 ATP (POTENTIAL).  
 FT DOMAIN 659 681 ACTIN-BINDING (BY SIMILARITY).  
 FT DOMAIN 761 775 ACTIN-BINDING (BY SIMILARITY).  
 FT MOD\_RES 130 130 METHYLATION (TRI-) (POTENTIAL).  
 FT MOD\_RES 699 699 ALKYLATION (SH-1) (POTENTIAL).  
 FT MOD\_RES 709 709 ALKYLATION (SH-2) (POTENTIAL).  
 FT CONFICT 1131 1131 A -> T (IN REF. 2).  
 FT CONFICT 1139 1139 G -> L (IN REF. 2).  
 FT CONFICT 1158 1158 G -> V (IN REF. 2).  
 FT CONFICT 1163 1163 A -> T (IN REF. 2).  
 FT CONFICT 1286 1289 TGG -> ONOV (IN REF. 2).  
 FT CONFICT 1302 1303 VS -> ET (IN REF. 2).  
 FT CONFICT 1451 1451 R -> T (IN REF. 2).  
 FT CONFICT 1470 1470 E -> V (IN REF. 2).  
 FT CONFICT 1470 1470

FT CONFICT 1473 1474 AS -> SF (IN REF. 2).  
 FT CONFICT 1569 1569 L -> V (IN REF. 2).  
 FT CONFICT 1598 1598 V -> E (IN REF. 2).  
 FT CONFICT 1606 1606 D -> N (IN REF. 2).  
 FT CONFICT 1643 1643 A -> D (IN REF. 2).  
 FT CONFICT 1648 1648 R -> Q (IN REF. 2).  
 FT CONFICT 1750 1750 Q -> K (IN REF. 2).  
 FT CONFICT 1822 1822 R -> K (IN REF. 2).  
 FT CONFICT 1845 1845 R -> H (IN REF. 2).  
 FT SEQUENCE 1939 AA; 223114 MW; 39ADB26AB79DFA53 CRC64;  
 SO  
 Query Match 14.0%; Score 149; DB 1; Length 1939;  
 Best Local Similarity 22.9%; Pred. No. 0.39;  
 Matches 52; Conservative 50; Mismatches 81; Indels 44; Gaps 7;  
 OY 24 LKNEIDNVAOLSOQDKKRDOSVITIDRLTLEPRNATVVSLOALKAEMLCSTLKKQ 83  
 Db 1207 LGEQIDNLRVQKLEKSEKMEITDLASNME-----TVSKAKGNLEKMCALDDQ 1259  
 OY 84 MKYLEQOODETR-----QOQEFAGRL-----RSKMKTMEQIEEL 117  
 Db 1260 LSEITKEEDQRLINDTAQPARLQTESGEYSRQDEKDLVLSQSGKAFTQIEEL 1319  
 OY 118 LOSQLEVEEMIRDMGVQSAVEQLAVYCVSLKKEYENLKEARKASGEVADKLKRDLEFS 177  
 Db 1320 KR-----QLEEEIKAKSALHALQSSRHQCDLRLQEQDEGEAKRLQRAMSKANSEVAMQ 1375  
 OY 178 RSKLOT-----VYSELDQAKLEKSNQKDLQSNQDKKMSLKKTLTMQ 220  
 Db 1376 RRTYETDAIQRTELEKAKKL--AQR-LQDAEHEVAVMKACSL 1419  
 RESULT 28  
 RA50\_ARCFU STANDARD; PRT; 886 AA.  
 ID RA50\_ARCFU  
 AC O29230;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE DNA double-strand break repair rad50 atpase.  
 GN RAD50 OR AF1032.  
 OS Archaeoglobus fulgidus.  
 CC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;  
 CC Archaeoglobus.  
 CC NCBI\_TaxID=2234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE=98049343; PubMed=9389475;  
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyprides N.C.,  
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uutterback T.,  
 RA Cotton M.D., Spriggs T., Artlisch P., Kaine B.P., Sykes S.M.,  
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the hyperthermophilic, sulphate-  
 reducing archaeon Archaeoglobus fulgidus.";  
 RL Nature 390:364-370(1997).  
 CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR). The  
 rad50/mre11 complex possesses single-strand endonuclease activity  
 and ATP-dependent double-strand-specific exonuclease activity.  
 CC Rad50 provides an ATP-dependent control of mre11 by unwinding  
 and/or prepositioning DNA ends into the mre11 active site (by  
 similarity).  
 CC -1- SUBUNIT: Forms a complex with mre11 (by similarity).  
 CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AE01032; AAB90211.1; -  
 DR TIGR: Afl032; -  
 DR InterPro: IPR003439; ABC\_transport.  
 DR InterPro: IPR001238; RecF.  
 DR Pfam: PF00470; RecF.1.  
 KM DNA repair: Hydrolase: ATP-binding: Coiled coil: Complete proteome.  
 FT NP\_BIND 31 38 ATP (BY SIMILARITY)  
 FT DOMAIN 148 728 COILED COIL (POTENTIAL).  
 SQ SEQUENCE 886 AA; 103633 MW; D35641D499AAB58 CRC64;

Query Match 13.9%; Score 148.5; DB 1; Length 886;  
 Best Local Similarity 23.0%; Pred. No. 0.19; Indels 45; Gaps 10;  
 Matches 59; Conservative 67; Mismatches 85;

QY 4 INKLFPLAEEENVDR---EFLKNELDNVRQA-----LSQDKE-KRDSQVI 48  
 DB 206 IESEKRELSSEVRLNELEKLEHNSRLSELRKQESSVLQVAGLEKLEKLEKQLEK 265  
 QY 49 IDTLRDLLEERNATVSLQ---QALGKAEKICSTLKQOMKYLEQO-----QDET 94  
 DB 266 VERLED-LEKKAKVEKLEKPKAEKRYSTLEKLEINQALRVEKREDLTFEAGIAQOL 324  
 QY 95 KOAGEAGRSKKKKTMEQIE-----LILQSLPEVEEKIRMGVQSGAVEQLAY 144  
 DB 325 KKAEDNSKLEETKRIEELRELEFEKSRHLETLKPKMDR---QGI-KATLEENL 380  
 QY 145 YCVSLKREYENLEKARKASGEVADLRKDLFSSRSKLTQTVYSELDOAKLELKSQKDIOS 204  
 DB 381 TPDKVEKMYDLSSKAKEEKEITEKLEK-LIAKKSSLKTRBAQKKAVEELKSAERCPV 439  
 QY 205 ADKEINSLKKLTMQ 220  
 DB 440 CGRELDEHHRKNIMAE 455

RESULT 29  
 KINH\_HUMAN  
 ID KINH\_HUMAN STANDARD; PRT; 963 AA.  
 AC P3176;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Kinesin heavy chain (ubiquitous kinesin heavy chain) (UKHC).  
 GN KIF5B OR KNS1 OR KNS.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta.  
 RX MEDLINE=92299683; PubMed=1607388;  
 RA Navone F., Nicolas J., Hom-Booher N., Sparks L., Bernstein H.D.,  
 RA McCaffrey G., Vale R.D.;  
 RT Cloning and expression of a human kinesin heavy chain gene:  
 RT interaction of the COOH-terminal domain with cytoplasmic microtubules  
 RT in transfected CV-1 cells.";  
 RL J. Cell Biol. 117:1263-1275(1992).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 1-349.  
 RX MEDLINE=96195066; PubMed=8606779;  
 RA Kull F.J., Sablin E.P., Lau R., Fletcher R.J., Vale R.D.;  
 RT Crystal structure of the kinesin motor domain reveals a structural

RT similarly to myosin.";  
 RL Nature 380:550-555(1996).  
 RN [3]  
 RP TISSUE SPECIFICITY.  
 RX MEDLINE=94242426; PubMed=7514426;  
 RA Nicolas J., Navone F., Hom-Booher N., Vale R.D.;  
 RT Cloning and localization of a conventional kinesin motor expressed  
 RT exclusively in neurons.";  
 RL Neuron 12:1059-1072(1994).

CC -1- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING  
 CC PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT.  
 CC -1- SUBUNIT: OLIGOMER COMPOSED OF TWO HEAVY CHAINS AND TWO LIGHT  
 CC CHAINS.  
 CC -1- SUBCELLULAR LOCATION: UNIFORMLY DISTRIBUTED BETWEEN THE CELL BODY  
 CC AND THE PROCESSES IN THE NEURONS.  
 CC -1- TISSUE SPECIFICITY: FOUND IN NERVEN AND ADULT BRAIN, LIVER,  
 CC KIDNEY, SPLEEN, HEART, LUNG AND SCIATIC NERVE.  
 CC -1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS: A LARGE GLOBULAR N-  
 CC TERMINAL DOMAIN WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY OF  
 CC KINESIN (IT HYDROLYSES ATP AND BINDS MICROTUBULE), A CENTRAL  
 CC ALPHA-HELICAL COILED COIL DOMAIN THAT MEDIATES THE HEAVY CHAIN  
 CC DIMERIZATION, AND A SMALL GLOBULAR C-TERMINAL DOMAIN WHICH  
 CC INTERACTS WITH OTHER PROTEINS (SUCH AS THE KINESIN LIGHT CHAINS),  
 CC VESICLES AND MEMBRANOUS ORGANELLES.  
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN  
 CC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X65873; CAA46703.1; -  
 DR PIR: A41919; A41919.  
 DR PDB: 1BG2; 14-OCT-98.  
 DR MIM: 602809; -  
 DR InterPro: IPR001752; kinesin.  
 DR Pfam: Pf00225; kinesin.  
 DR PRINTS: PR00380; KINESINHEAVY.  
 DR SMART: SM00129; KISC; 1.  
 DR PROSITE: PS00411; KINESIN\_MOTOR\_DOMAIN; 1.  
 DR PROSITE: PS50067; KINESIN\_MOTOR\_DOMAIN; 1.  
 KM Motor protein; Microtubules; ATP-binding; Coiled coil; 3D-structure.  
 FT DOMAIN 1 256 KINESIN-MOTOR.  
 FT DOMAIN 329 914 KINESIN-MOTOR.  
 FT DOMAIN 915 963 COILED COIL.  
 FT NP\_BIND 85 92 GLOBULAR.  
 FT NP\_BIND 85 92 ATP (BY SIMILARITY).  
 SQ SEQUENCE 963 AA; 109685 MW; A1P5760C3250C8B CRC64;

Query Match 13.9%; Score 148.5; DB 1; Length 963;  
 Best Local Similarity 21.5%; Pred. No. 0.21; Indels 95; Gaps 9;  
 Matches 67; Conservative 55; Mismatches 94;

QY 4 INKLFPLAEEENVDR---IDREFL---KNELDNVRQIS----- 36  
 DB 425 IAKLYKQDDDERINQOSQLVEKLTQMDQEBLLASTRRQDNMAELNRLQDAENDAS 484  
 QY 37 -----QKDKERDSQVIIDTLRDLLEERNATVVSLSQALCK----- 72  
 DB 485 KEVKEVLAQLEELAVNYDKSDEVEKTYEYELLSLQKSTSLASISAELOKLEMT 544  
 QY 73 -----AEMLCSTLK-----KOMYLVQOQDETQQAQDEAGRLSKKMTM 111  
 DB 545 NHQKRAAEKMSLTKLAETIGIAGVGNNDYKQPGSTGMIDEFYVALIYLSKMSSEVKTM 604  
 QY 112 -----EQLELLQSLPEVEEKIRMGVQSGAVEQLAYCVSLKREYNLEKARRASGEVA 167  
 DB 605 VKRCKQLESQTESNKKMEENEKELACQRLISQHEAKIKISITREYLVNVEQKKRQLEESV 664

```

OY 168 DKLRLDFFSR-----SKLOTVYSELDQA-----KLELKSQKDLQSGADKEI 209
Db 665 DALSEELVOLRAQEKVHEMEKEHLNKVOTA-NEVKAQAVEQIQSHRETHQKQISSLRDEY 723
OY 210 MSLKKKLTMLQ 220
Db 724 EAKAKLTLDLQ 734

RESULT 30
MYH7_MESAU STANDARD: PRT: 1934 AA.
AC P13540; Q60540;
DT 01-JAN-1990 (Rel. 13, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE MYOSIN heavy chain, cardiac muscle beta isoform (MYHC-beta).
GN MYH7.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FLB; TISSUE=Liver;
RX MEDLINE=95115033; PubMed=7815459;
RA Wang R., Sole M.J., Cukerman E., Liew C.-C.;
RT "Characterization and nucleotide sequence of the cardiac alpha-myosin
RT heavy chain gene from Syrian hamster.";
RL J. Mol. Cell. Cardiol. 26:1155-1165(1994).
RN [2]
RP SEQUENCE OF 962-1935 FROM N.A.
RX MEDLINE=88247788; PubMed=3380703;
RA Jandreski M.A., Sole M.J., Liew C.-C.;
RT "Sequence of cDNA encoding the Syrian hamster cardiac beta-myosin
RT heavy chain.";
RL Nucleic Acids Res. 16:4737-4737(1988).
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (SL) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L12104; AAA62313.1; -
DR EMBL; X07273; CAA30256.1; -
DR PIR; A28298; A28298.
DR HSSP; P08799; IMMD.
DR InterPro; IPR000048; IO.
DR InterPro; IPR004003; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; myosin_N; 1.
DR Pfam; PF01576; myosin_tail; 1.

```

```

DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Methylation; Alkylation; Multigene family;
KW Calmodulin-binding.
FT DOMAIN 1 779 MYOSIN HEAD-LIKE.
FT DOMAIN 780 809 COILED COIL (POTENTIAL).
FT DOMAIN 839 1934 ATP.
FT NP_BIND 177 184 ACTIN-BINDING.
FT DOMAIN 654 676 ACTIN-BINDING.
FT DOMAIN 756 770 ACTIN-BINDING.
FT MOD_RES 128 128 METHYLATION (TR1-) (POTENTIAL).
FT MOD_RES 694 694 ALKYLATION (SH-1) (POTENTIAL).
FT MOD_RES 704 704 ALKYLATION (SH-2) (POTENTIAL).
FT MOD_RES 966 966 D->E (IN REF. 2).
FT CONFLICT 966 966 T->TE (IN REF. 2).
FT CONFLICT 978 978 D->Q (IN REF. 2).
FT CONFLICT 986 986 D->Y (IN REF. 2).
FT CONFLICT 1008 1014 D->V (IN REF. 2).
FT CONFLICT 1057 1057 D->N (IN REF. 2).
FT CONFLICT 1057 1057 D->D (IN REF. 2).
FT CONFLICT 1095 1095 D->N (IN REF. 2).
FT CONFLICT 1217 1217 D->D (IN REF. 2).
FT CONFLICT 1271 1271 T->A (IN REF. 2).
FT CONFLICT 1327 1327 C->R (IN REF. 2).
FT CONFLICT 1358 1358 L->V (IN REF. 2).
FT CONFLICT 1504 1504 M->L (IN REF. 2).
FT CONFLICT 1537 1537 N->K (IN REF. 2).
FT CONFLICT 1556 1556
SQ SEQUENCE 1934 AA; 222928 MW; FDBAC58310B0B57D CRC64;

```

```

Query Match 13.9%; Score 148.5; DB 1; Length 1934;
Best Local Similarity 20.5%; Pred. No. 0.42;
Matches 61; Conservative 60; Mismatches 88; Indels 89; Gaps 9;

```

```

OY 1 RTIINKLFFDLAEENLVDR-----EFLK----- 25
Db 1128 RAKVEKRLSDLSRELEFSEIRLEFAGATSVQIEKNKKREAFQKMRDLEATLQHEAT 1187
OY 26 -----NELDNVRAQLSQDKERKDSQVYITDLRDTLEARNATVVSLOQA 69
Db 1188 AAALRRKHADSVALGEQIDNLQRYKQLEKESEFKELDDVYSNMGO-----LIKA 1240
OY 70 LGKAEMLCSTLAKQKMTYLEQOODEFKQAQEFAGRLRSKMT-----MQEILLDSQ 121
Db 1241 KANLEKMCRTLEDQNEHRSKAEETQRSVNDLTQSRAKLTQENGLSRQDKEKELI-SQ 1299
OY 122 LP-----EVEEMIRDMGVGQSAVEQLAVYCVSLKKEYENLKEARKASGEV 166
Db 1300 LTRGKLTYYTOQLDLKRLQEEVEVAKAKNTLAHLASARDCDLLRQYEEFEAKAELOCV 1359
OY 167 ADKLRLDFFSRSLQQT---VYSELDQAKLELSAQKDLQSGADKEIMSLKKLTMLQ 220
Db 1360 LSKANSEVAQRTYETDAIQRTLEBEAKKKL--AQK-LQDAEBAVAVNAKSSLE 1414

RESULT 31
MYSS_CHICK STANDARD: PRT: 1938 AA.
AC P13538; O13228;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, skeletal muscle, adult.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]

```



RP SEQUENCE FROM N.A.  
RC STRAIN-WHITE LECHORN; TISSUE-Pectoralis muscle;  
RA Chao T.H., Bandman E., Moore L.;  
RT "Cloning, nucleotide sequence and characterization of a full-length  
RT cDNA encoding the myosin heavy chain from adult chicken pectoralis  
RT major muscle.";  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
RN [12]  
RP SEQUENCE OF 1-205.  
RC TISSUE-Pectoralis muscle;  
RX MEDLINE=92041767; PubMed=1939027;  
RA Hayashida M., Maiba T., Matsuda G.;  
RT "The primary structure of skeletal muscle myosin heavy chain: I.  
RT Sequence of the amino-terminal 23 kDa fragment.";  
RL J. Biochem. 110:54-59(1991).  
RN [13]  
RP SEQUENCE OF 206-636.  
RC TISSUE-Pectoralis muscle;  
RX MEDLINE=92041768; PubMed=1939028;  
RA Komine Y., Maiba T., Matsuda G.;  
RT "The primary structure of skeletal muscle myosin heavy chain: II.  
RT Sequence of the 50 kDa fragment of subfragment-1.";  
RL J. Biochem. 110:60-67(1991).  
RN [14]  
RP SEQUENCE OF 637-837.  
RC TISSUE-Pectoralis muscle;  
RX MEDLINE=92041769; PubMed=1939029;  
RA Maiba T., Miyaniishi T., Matsuzono K., Tanioaka Y., Matsuda G.;  
RT "The primary structure of skeletal muscle myosin heavy chain: III.  
RT Sequence of the 22 kDa fragment and the alignment of the 23 kDa, 50  
RT kDa, and 22 kDa fragments.";  
RL J. Biochem. 110:68-74(1991).  
RN [15]  
RP SEQUENCE OF 838-1938.  
RC TISSUE-Pectoralis muscle;  
RX MEDLINE=92041770; PubMed=1939030;  
RA Maiba T., Yajima E., Nagata S., Miyaniishi T., Nakayama S., Matsuda G.;  
RT "The primary structure of skeletal muscle myosin heavy chain: IV.  
RT Sequence of the rod, and the complete 1,938-residue sequence of the  
RT heavy chain.";  
RL J. Biochem. 110:75-87(1991).  
RN [16]  
RP PRELIMINARY SEQUENCE OF 1-808.  
RX MEDLINE=87092420; PubMed=3467365;  
RA Maiba T., Hayashida M., Tanioaka Y., Komine Y., Matsuda G.;  
RT "The primary structure of the myosin head.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:416-420(1987).  
RN [17]  
RP SEQUENCE OF 842-1270.  
RX MEDLINE=90121764; PubMed=2610940;  
RA Maiba B.;  
RT "Complete amino-acid sequence of subfragment-2 in adult chicken  
RT skeletal muscle myosin.";  
RL Biol. Chem. Hoppe-Seyler 370:1027-1034(1989).  
RN [18]  
RP SEQUENCE OF 852-1108.  
RX MEDLINE=89374803; PubMed=2775482;  
RA Maiba B.;  
RT "Amino-acid sequence of the short subfragment-2 in adult chicken  
RT skeletal muscle myosin.";  
RL Biol. Chem. Hoppe-Seyler 370:549-558(1989).  
RN [19]  
RP SEQUENCE OF 1145-1270.  
RX MEDLINE=89228549; PubMed=2713098;  
RA Maiba B.;  
RT "Amino-acid sequence of the hinge region in chicken myosin  
RT subfragment-2.";  
RL Biol. Chem. Hoppe-Seyler 370:55-61(1989).  
RN [10]  
RP SEQUENCE OF 1857-1938 FROM N.A.  
RX MEDLINE=87217964; PubMed=5034534;  
RA Moriarty D.M., Barringer K.J., Dodgson J.B., Richter H.E.,  
RA Young R.B.;

RT "Genomic clones encoding chicken myosin heavy-chain genes.";  
RL DNA 6:91-99(1987).  
RN [11]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 1-843.  
RX MEDLINE=93303624; PubMed=8316857;  
RA Rayment I., Rypniewski W.R., Schmidt-Bae K., Smith R.,  
RA Tomchick D.R., Bening M.M., Winkelmann D.A., Wesenberg G.,  
RA Holden H.M.;  
RT "Three-dimensional structure of myosin subfragment-1: a molecular  
RT motor.";  
RL Science 261:50-58(1993).  
CC -1- FUNCTION: MUSCLE CONTRACTION. MYOSIN IS A PROTEIN THAT BINDS TO  
CC F-ACTIN AND HAS ATPASE ACTIVITY THAT IS ACTIVATED BY F-ACTIN.  
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE  
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
CC SUBFRAGMENT (S2).  
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: U87231; AAB47555.1; -;  
DR EMBL: M16557; AAA48970.1; -;  
DR PIR: PX0050; PX0051.  
DR PIR: A26821; A26821.  
DR PIR: S02082; S02082.  
DR PIR: S04501; S04501.  
DR PIR: S05515; S05515.  
DR PDB: 2MYS; 11-JAN-97.  
DR InterPro: IPR000048; IQ.  
DR InterPro: IPR004009; Myosin\_N.  
DR InterPro: IPR002928; Myosin\_tail.  
DR InterPro: IPR001609; Myosin\_head.  
DR Pfam: PF00612; IQ; 2.  
DR Pfam: PF00063; myosin\_head; 1.  
DR Pfam: PF02736; Myosin\_N; 1.  
DR Pfam: PF01576; Myosin\_tail; 1.  
DR ProDom: PD000355; myosin\_head; 1.  
DR SMART: SM00015; IQ; 1.  
DR SMART: SM00242; MYSC; 1.  
DR PROSITE: PSS0096; IQ; 1.  
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
KW ATP-binding; Methylation; Alkylation; Phosphorylation; Acetylation;  
KW Calmodulin-binding; Multigene family; 3D-structure.  
FT INT\_MET 0  
FT DOMAIN 1 782 812 MYOSIN HEAD-LIKE.  
FT DOMAIN 783 838 IQ.  
FT DOMAIN 838 840 HINGE.  
FT DOMAIN 841 1938 COILED COIL (POTENTIAL).  
FT NP\_BIND 179 186 ATP (POTENTIAL).  
FT DOMAIN 657 679 ACTIN-BINDING.  
FT DOMAIN 773 773 ACTIN-BINDING.  
FT MOD\_RES 1 1 ACETYLATION.  
FT MOD\_RES 35 35 METHYLATION (MONO-).  
FT MOD\_RES 130 130 METHYLATION (TRI-).  
FT MOD\_RES 551 551 METHYLATION (TRI-).  
FT MOD\_RES 755 755 METHYLATION (MONO-).  
FT MOD\_RES 755

```

FT MOD_RES 697 697 ALKYLATION (SH-1).
FT MOD_RES 707 707 ALKYLATION (SH-2).
FT CONFLICT 907 907 C -> Q (IN REF. 7 AND 8).
FT CONFLICT 980 980 L -> F (IN REF. 1).
FT CONFLICT 1343 1343 E -> D (IN REF. 5).
FT CONFLICT 1545 1545 S -> A (IN REF. 5).
FT CONFLICT 1796 1797 HV -> QL (IN REF. 5).
FT CONFLICT 1830 1830 S -> A (IN REF. 5).
FT CONFLICT 1863 1863 I -> V (IN REF. 10).
FT CONFLICT 1929 1931 IHG -> FH (IN REF. 10).
SO SEQUENCE 1938 AA; 223013 MW; EDD01CBA2681E10F CRC64;

Query Match 13.9%; Score 148.5; DB 1; Length 1938;
Best Local Similarity 22.6%; Pred. No. 0.42;
Matches 53; Conservative 50; Mismatches 80; Indels 51; Gaps 7;

QY 24 LKNELDNRAQLSOKDKRDSQVYIIDLRTLEERNATVVSLOQALKAEMLCSTLKQ 83
DQ 1205 LGEQIDNLRVKKQKLEKSELMKMDLSDASNM-----SVSKAKNLEKMCRTLEDQ 1257
QY 84 MKYLEQOODE-----TKQAEAGRL-----RSKMKTMQIEEL 117
DQ 1258 LSEIKTEEQNRMINDTQARLQETGYSKQAEKDALISLQSGKQFTQOIEEL 1317
QY 118 LQSLQPEVEEMIRDMGVQSAVEQLAVYCVSLKKEYENLKEAR-----KASGEVAD- 168
DQ 1318 KR---HLEEFKAKNALNALQSAHRDCELLREQYEEQEKGLQALSKANSEVAVQW 1373
QY 169 --KLKRDFFSRSLQTYVSELDQAKLEKSAQKDLQSADEKINSKRLKMLQ 220
DQ 1374 RTKYTEYDRIQRTLEELKRLAQ---RLQDAEHVEHVAVNAKCASTLEKQRLQ 1424

RESULT 32
MYH9_CHICK STANDARD; PRT; 1959 AA.
AC P14105;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, nonmuscle (Cellular myosin heavy chain) (NMHC).
GN MYH9.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestinal epithelium;
RX MEDLINE=90046668; PubMed=2813355;
RA Shohet R.V., Conti M.A., Kawamoto S., Preston Y.A., Brill D.A.,
RA Adelstein R.S.;
RT "Cloning of the cDNA encoding the myosin heavy chain of a vertebrate
RT cellular myosin.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:7726-7730(1989).
CC -1- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS,
CC CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND
CC CAPING.
CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY
CC CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2
CC REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M26510; AAA48974.1; -.
CC PIR: A33877; A33977.
CC HSSP: P10587; IBL1.
CC InterPro: IPR000048; IO.
CC InterPro: IPR004009; Myosin_N.
CC InterPro: IPR002928; Myosin_tail.
CC InterPro: IPR001609; myosin_head.
CC Pfam: PF00612; IO; 1.
CC Pfam: PF00063; myosin_head; 1.
CC Pfam: PF02736; Myosin_N; 1.
CC Pfam: PF01576; Myosin_tail; 1.
CC PRINTS: PR00193; MYOSINHEAVY.
CC PRODOM: PD000355; myosin_head; 1.
CC SMART: SM00015; IO; 1.
CC SMART: SM00242; MYSC; 1.
CC PROSITE: PS50096; IO; 1.
CC MYOSIN; ATP-binding; Calmodulin-binding; Actin-binding;
CC Coiled coil; Alkylation; Multigene family.
CC MYOSIN HEAD-LIKE.
CC DOMAIN 1 778
CC DOMAIN 779 808
CC NP_BIND 837 1925 COILED COIL (POTENTIAL).
CC FT 174 181 ATP (POTENTIAL).
CC FT 654 676 ACTIN-BINDING.
CC FT MOD_RES 694 694 ALKYLATION (SH-1) (POTENTIAL).
CC FT MOD_RES 704 704 ALKYLATION (SH-2) (POTENTIAL).
CC SO SEQUENCE 1959 AA; 226502 MW; A75C86086FD3A1A1 CRC64;

```

```

Query Match 13.9%; Score 148.5; DB 1; Length 1959;
Best Local Similarity 22.7%; Pred. No. 0.42;
Matches 53; Conservative 56; Mismatches 95; Indels 29; Gaps 8;

QY 11 LAQEEVYLDREFIKNELDNYRAQLSOKDKERDSQVYIIDLND-----TLEER 59
DQ 1665 LAQAKENKRLKLSMEAEWIOLEFIAAEBAKRAQOQERDELDAINSSGKALAMEK 1724
QY 60 ---NATVVSLOQAL-----GRAEMLCSTLKKQMKYLBQOQDITKQAEAGLRKMKTM 112
DQ 1725 RRLERARINQLEELKEEGNTEIINDRKANLQIDMNAVLNERNNAQKNENARQME 1784
QY 113 QTELLSOLPEVEEMIRDMGVQSAVEQLAVYCVSLKKEYE-NLKPAKASGEV--ADK 169
DQ 1785 RQNELKIKLOEMESAYKSK--YKATITALEAKTYQLEQDMDMTKERQASKQVRAEK 1842
QY 170 LRKDLF---SSRSKLQTYVSELDQAKLEKSAQKDLQSADEKIM--SLKRL 216
DQ 1843 KLKQIDLLQVDDERRNNAQFKQADKAMWRLKQLKRLQEEAEAEQARNAVRRKL 1895

RESULT 33
MYSB_CAEEL STANDARD; PRT; 1966 AA.
AC P02566;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain B (MHC B).
GN UNC-54 OR MYO-4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83273600; PubMed=6576334;
RA Karn J., Brenner S., Barnett L.;
RA "Protein structural domains in the Caenorhabditis elegans unc-54
RT myosin heavy chain gene are not separated by introns.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).

```



DR Pfam; PF00612; IQ; 1.  
 DR Pfam; PF00063; myosin\_head; 1.  
 DR Pfam; PF02736; Myosin\_N; 1.  
 DR Pfam; PF01576; Myosin\_tail; 1.  
 DR PRINTS; PR00193; MYOSINHEAVY.  
 DR Pfam; PF00035; myosin\_head; 1.  
 DR SMART; SM0015; IQ; 1.  
 DR SMART; SM00242; MYSC; 1.  
 DR PROSITE; PS50096; IQ; 1.  
 DR Myosin: Alternative splicing: Coiled coil; Actin-binding;  
 KW ATP-binding; Calmodulin-binding.  
 FT DOMAIN 1 829  
 FT MYOSIN HEAD-LIKE.  
 FT DOMAIN 830 859  
 FT IQ.  
 FT NP\_BIND 886 2017  
 FT COILED COIL (POTENTIAL).  
 FT NP\_BIND 225 232  
 FT ATP.  
 FT DOMAIN 250 260  
 FT 25 KDA/50 KDA JUNCTION.  
 FT DOMAIN 682 694  
 FT 50 KDA/20 KDA JUNCTION.  
 FT DOMAIN 705 727  
 FT ACTIN-BINDING.  
 FT DOMAIN 742 758  
 FT REACTIVE SULFHYDRYL/ACTIN-BINDING.  
 FT DOMAIN 1303 2017  
 FT LIGHT MEROMYOSIN (LMW).  
 FT DOMAIN 1303 1970  
 FT ALPHA-HELICAL TAILPIECE (LMW).  
 FT DOMAIN 1971 2017  
 FT GLOBULAR TAILPIECE.  
 FT VARSPLIC 1 45  
 FT MISSING (IN SHORT ISOFORM).  
 SQ SEQUENCE 2017 AA: 232016 MW: 73E3CB02BA8F2528 CRC64;

Query Match 13.8%; Score 147.5; DB 1; Length 2017;  
 Best Local Similarity 23.7%; Pred. No. 0.49; Indels 33; Gaps 7;  
 Matches 57; Conservative 49; Mismatches 102;

QY 1 RTIIKLEFDLAEEENVLDRFLKNELDNVAQLSQDKER--DSQVI-----ID 50  
 Db 1263 KTVLEKAKGTLEAENAD-----LATELRSVNSSRQENDRRKKAESQIAELQYVLAIE 1316  
 QY 51 TLRLTLEARNATVVSLOQ-----ALGKAEMLCSTLKKQKMYLEQODETKQAOEENG 103  
 Db 1317 RARSELQER---CTKLOQEAENITNQLTEAEIKSAVAKSSANMSQSLTEAOQLLEETR 1373  
 QY 104 ---LRSMKMTMEQIELLOSLPEVEEMIRDMGVSQSAVOLAVYCVSLKEVENLKEA 159  
 Db 1374 QKGLSSKRLROISEKELQLEDEDDAKKNY---ERKLEVTITQOMEIKKKAEEEDDL 1430  
 QY 160 RKASGEVADKLKDLFSSRSKLTQTVYSELDAQLELSAQKDLQADKESLKKLTML 219  
 Db 1431 AKLEEEGKKRLKNDIEALERQVKELIAQNDRDKSKKIKQSELEDDATELEAQRKVL 1490  
 QY 220 Q 220  
 Db 1491 E 1491

RESULT 35  
 CENE\_HUMAN STANDARD; PRT: 2663 AA.  
 AC 002224;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Centromeric protein E (CENP-E protein).  
 GN CENPE.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93024922; Pubmed=1406971;  
 RA Yen T.J., Li G., Schaar B.T., Szilak I., Cleveland D.W.;  
 RT "CENP-E is a putative kinetochore motor that accumulates just before  
 mitosis";  
 RL Nature 359:536-539(1992).  
 RN [2]  
 RP CHARACTERIZATION.

RX MEDLINE=95196755; Pubmed=7889940;  
 RA Thrower D.A., Jordan M.A., Schaar B.T., Yen T.J., Wilson L.;  
 RT "Mitotic HeLa cells contain a CENP-E-associated minus end-directed  
 microtubule motor";  
 RL EMBO J. 14:918-926(1995).  
 RN [3]  
 RP CHARACTERIZATION.  
 RX MEDLINE=98437347; Pubmed=9763420;  
 RA Chan G.K.T., Schaar B.T., Yen T.J.;  
 RT "Characterization of the kinetochore binding domain of CENP-E reveals  
 interactions with the kinetochore proteins CENP-F and hNubrl";  
 RL J. Cell Biol. 143:49-63(1998).  
 CC -1- FUNCTION: MINUS-END DIRECTED MICROTUBULE MOTOR. PROBABLE  
 KINETOCHORE MOTOR. ACCUMULATES JUST BEFORE MITOSIS AT THE G2 PHASE  
 OF THE CELL CYCLE. PROBABLY IMPORTANT FOR CHROMOSOME MOVEMENT  
 AND/OR SPINDLE ELONGATION.  
 CC -1- SUBUNIT: INTERACTS WITH CENP-F AND BUBR1 KINASE.  
 CC -1- SUBCELLULAR LOCATION: ASSOCIATES WITH KINETOCHORES DURING  
 CONGRESSION, RELOCATES TO THE SPINDLE MIDZONE AT ANAPHASE, AND IS  
 QUANTITATIVELY DISCARDED AT THE END OF THE CELL DIVISION.  
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z15005; CAAT8727.1; -.  
 DR PIR; S28261; S28261.  
 DR HSSP; P17119; 3KAR.  
 DR MIM; 117143; -.  
 DR InterPro; IPR001752; kinesin.  
 DR Pfam; PF00225; kinesin.  
 DR PRINTS; PR00380; KINESINHEAVY.  
 DR SMART; SM00129; KISC; 1.  
 DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN; 1.  
 DR PROSITE; PS50067; KINESIN\_MOTOR\_DOMAIN; 1.  
 KW Motor protein; Cell division; ATP-binding; Coiled coil; Mitosis;  
 KM Cell cycle; Centromere.  
 FT DOMAIN 1 335  
 FT KINESIN-MOTOR.  
 FT DOMAIN 336 2471  
 FT COILED COIL (POTENTIAL).  
 FT DOMAIN 2472 2663  
 FT GLOBULAR (POTENTIAL).  
 FT NP\_BIND 86 93  
 FT ATP (BY SIMILARITY).  
 SQ SEQUENCE 2663 AA: 312087 MW: CEFCl3880C8C8CB8 CRC64;

Query Match 13.8%; Score 147.5; DB 1; Length 2663;  
 Best Local Similarity 25.1%; Pred. No. 0.65;  
 Matches 57; Conservative 48; Mismatches 91; Indels 31; Gaps 9;  
 QY 13 QEEENVLDREFLKNELDNVAQLSQDKER-RDSQVIITLRLTLEARNATVVSLOALG 71  
 Db 1638 ETOEKMEIEHLKEQFOFKLNIENIETIRLQI---LHENLEEMR-SVTKERDRLR 1692  
 QY 72 KAEMLCSTLKKQMYLEQODET---KQAOEENG---RLRSKMTMEQIELLOSLP 123  
 Db 1693 SVE---ETLKVERDOLKENLRETTITRDLEKQDELKIYMHMLKEHOETIDKLRGIVSEKTN 1749  
 QY 124 EVEEMIRDMGVSQSAVOLAVYCVSLKEVENLKEARKSGEYADKL---RK 172  
 Db 1750 EISMNQKDLHNSNALKAQDLKIOBELRLIAHMLKEQOETIDKLRGIVSEKTDKLSMNQK 1809  
 QY 173 DLFSRSKLTQTVYSELDAQLELSAQKDLQADK---EIMSLKKL 216  
 Db 1810 DLHNSNAKLOEKIOELKANHQTLTKKDVNETQKVSSEMQLKKQI 1856

RESULT 36  
 Y373\_BOVIN STANDARD; PRT: 1453 AA.  
 ID Y373\_BOVIN

AC Q9TU23;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein KIAA0373.  
GN KIAA0373.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovine; Bos.  
OX NCBI\_TaxID=9913;  
RN  
RP SEQUENCE FROM N.A.  
RC TISSUE=Trachea;  
RA Jovov B., Ripoll P.J., Benos D.J.;  
RT Submitted (MIG-1999) to the EMBL/GenBank/DBJ databases.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL; AF176816; AAF00990.1; .  
KM Hypothetical protein: Coiled coil.  
FT DOMAIN 37 1426  
SQ SEQUENCE 1453 AA; 169934 MW; 03CBA02A64CF4139 CRC64;

Query Match 13.8%; Score 147; DB 1; Length 1453;  
Best Local Similarity 23.1%; Pred. No. 0.37; Mismatches 86; Gaps 11;  
Matches 64; Conservative 49; Indels 86; Gaps 11;  
QY 15 EENVLDREFLKNEI-----DNVRAQLSQKDKERDSQVYIIDLTLDLLEERNATV 64  
DB 571 EENILSRDKVINEIRLRPATAEQEKLAEPKREVEPKS-----HHTLKLAHGTIA 622  
QY 65 SLQALGKAEMLCTLKQKMYLEQODE-----TKQAEELGRLSKKKTMEQ----- 113  
DB 623 NMQARLNQKE---EVLKRYOHLLEKAREQREIVKHEEELHTLRKLELDQADNSLKF 679  
QY 114 ---TELLLOS--OLPEVEEMIRDMGVGSAVEQ---LAVYCVSLK----- 150  
DB 680 ETAMDILKQSTPYPTKHFIRLAEMETVAEODDSLSLVKIKQVSQDLEROKETTEL 739  
QY 151 --KEYENLK-----EARKASGEVADKLKDLFSSRSKLTQTVYSEL----- 188  
DB 740 KIKFEFENMKRLQENHADVKKIKAEVED--LRCLLVQSQKSSQSLKSELQTKREANSRAP 798  
QY 189 -----DQAKLELSAQKDLQSDAKKEIMSLKKLT 217  
DB 799 TTTMRNLVERLKSQALAEKQKALSRALLELRAMET 835

RESULT 37  
RA50\_THEVO  
ID RA50\_THEVO STANDARD; PRT; 895 AA.  
AC P58302;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE DNA double-strand break repair Rad50 ATPase.  
GN RAD50 OR TVG023531.  
OS Thermoplasma volcanium.  
OC Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmaceae;  
OC Thermoplasma.  
OX NCBI\_TaxID=50339;  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=GSI / DSM 4299 / JCM 9571;  
RX MEDLINE=20570466; PubMed=11121031;

RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,  
RA Kawashima Ohya Y., Matanabe K., Yamazaki M., Kanehori K., Kawamoto T.,  
RA Nishishita T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;  
RT "Archaeal adaptation to higher temperatures revealed by genomic  
RT sequence of Thermoplasma volcanium.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).  
CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR). The  
CC rad50/mre11 complex possesses single-strand endonuclease activity  
CC and ATP-dependent double-strand-specific exonuclease activity.  
CC Rad50 provides an ATP-dependent control of mre11 by unwinding  
CC and/or repositioning DNA ends into the mre11 active site (By  
CC similarity).  
CC -1- SUBUNIT: Forms a complex with mre11 (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL; AP000991; BAB59370.1; .  
DR InterPro: IPR003439; ABC\_transportr.  
DR InterPro: IPR000875; Cecropin.  
DR InterPro: IPR001238; Recf.  
DR InterPro: IPR003395; SMC\_N.  
DR Pfam: PF004470; Recf; 1.  
DR Pfam: PF02463; SMC\_N; 1.  
KM DNA repair: Hydrolase; ATP-binding; Coiled coil; Complete proteome.  
FT NP\_BIND 30 37  
FT DOMAIN 171 737  
SQ SEQUENCE 895 AA; 103223 MW; 4A47DA5287E82D3A CRC64;

Query Match 13.7%; Score 146.5; DB 1; Length 895;  
Best Local Similarity 22.8%; Pred. No. 0.24; Mismatches 82; Indels 53; Gaps 8;  
Matches 55; Conservative 51; Indels 53; Gaps 8;  
QY 14 EENVLDRE-----EFLKNELDNVRAQLSQKDKERDSQVYIIDLTLDLLEERNAT 62  
DB 150 ERKRLDELINLDRLGAYLLKLEVIDLT/ANVSDYDLKNELOSKI-----NEIDNNKQ 205  
QY 63 VVSLQALGKAEMLCTLKQKMYLEQODETKQAEAGLRKSKMTMEQIEL--LQSQ 121  
DB 206 IEELSKRLRLPEIKALEEINIKENKKD--HUNEELRLNQLLETIKKYEMLAESQ 262  
QY 122 -----LPEVEEMIRDMGVGSAVEQLAVYCVSLKKEYENLKARKASGEVADK 169  
DB 263 SRKASIEMEVYKLPSIEELKRLNNAVYKRNELI-----EYINKKDLGSLSEIES 316  
QY 170 LRKDLF---SSRSKLTQTVY-----EIDQAKLELSAQKDLQSDAKKEIMSLK 213  
DB 317 LKSDLKYEAKRKLELDQSPFSEFLKRRKRELDLKRSLKDEEDNYGSAVNIENIK 376  
QY 214 K 214  
DB 377 K 377

RESULT 38  
MYSS\_RABIT  
ID MYSS\_RABIT STANDARD; PRT; 1084 AA.  
AC P02562;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Myosin heavy chain, skeletal muscle (Fragments).  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;

1	SEQUENCE OF 1-258.
2	Capony J.-P., Elzinga M.;
3	"The amino acid sequence of A 34,000 dalton fragment from S-2 of
4	myosin.";
5	Biophys. J. 33:148A-148A(1981).
6	[2]
7	SEQUENCE OF 259-428.
8	MEDLINE=85131142; PubMed=3972832;
9	Lu R.C., Wong A.;
10	"The amino acid sequence and stability predictions of the hinge
11	region in myosin subfragment 2.";
12	J. Biol. Chem. 260:3456-3461(1985).
13	[3]
14	SEQUENCE OF 409-1084 FROM N.A.
15	MEDLINE=87304245; PubMed=3305014;
16	Maeda K., Sczakiel G., Wittinghofer A.;
17	"Characterization of cDNA coding for the complete light meromyosin
18	portion of a rabbit fast skeletal muscle myosin heavy chain.";
19	Eur. J. Biochem. 167:97-102(1987).
20	[4]
21	FUNCTION: MUSCLE CONTRACTION.
22	[5]
23	SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
24	HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
25	AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
26	[6]
27	SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
28	[7]
29	DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
30	CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
31	CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
32	[8]
33	MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
34	MEROMYOSIN (LMN) AND 1 HEAVY MEROMYOSIN (HNM). IT CAN LATER BE
35	SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
36	SUBFRAGMENT (S2).
37	[9]
38	SIMILARITY: THE PERIODICITIES OF HYDROPHOBIC AND CHARGED RESIDUES,
39	WHICH DICTATE THE ALPHA-HELICAL COILED-COIL STRUCTURE ARE
40	CONSERVED.
41	-----
42	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
43	BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
44	THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
45	USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
46	MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
47	ENTITIES REQUIRES A LICENSE AGREEMENT (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
48	OR SEND AN EMAIL TO <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
49	-----
50	EMBL; X03958; CAA29391.1; -.
51	PIR; A02985; A02985.
52	PIR; A05280; A05280.
53	PIR; S00084; S00084.
54	HSSP: P02649; INPN.
55	PFam: PF01576; Myosin_tail; 1.
56	DR Pfam: Muscle protein; Coiled coil; Thick filament; Actin-binding;
57	Kw Art-binding; Multigene family.
58	FT NON_TER 1 1
59	FT DOMAIN <1 258 ALPHA-HELICAL TAILPIECE (S2).
60	FT NON_CONS 258 259
61	FT DOMAIN <259 1084 RODLIKE TAIL (S2 AND LMN DOMAINS).
62	FT DOMAIN 435 1084 COILED COIL (POTENTIAL).
63	FT DOMAIN 405 405 L -> V.
64	FT VARIANT 408 408 V -> L.
65	FT VARIANT 421 421 E -> D.
66	FT VARIANT 423 423 S -> G.
67	FT VARIANT 426 426 K -> R.
68	SEQUENCE 1084 AA; 125488 MW; 229CFD69A6E1F7F0 CRC64;

```

Query Match      13.7%  Score 146;  DB 1;  Length 1084;
Best Local Similarity 24.5%  Pred. No. 0.32;
Matches 47;  Conservative 42;  Mismatches 83;  Indels 20;  Gaps 6;

OY 24 LKNELDNVAALQSQDKKEKRDQVITDRLTLEERNATVVSLOALGAENLCTLKKQ 83
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
352 LGGDIDINQARKKLEKSEIKMIEDIDLAGNME-----TVSKAKNLEKMCITLEQ 404

```

OY	84	MKYLEOODDFKQAOEKGRLRSKMKTMEQJELLLOSLPVEEIMIDMGVGSAYBOLA	143
	:	:::::	:
	:	::::	:
Db	405	LSEVKTKFEEHQRILNELISAOKARLHT--ESGE--FSROLEKDMMVSOLSRGGAFTQ--	459
OY	144	VYCVALKEYEENLKEARKASGEVADKLRLDKLFSSRSKIQTIVSELDOAKLELSAOKDLQ	203
	:	:::::	:
	:	::::	:
Db	460	-QIGSLRKOLEEPFKAKSALAHALOSSRDC---DILREQYEEDQAEL---QRAMS	511
OY	204	SADREIMSIRKK	215
	:	::::	:
Db	512	KANSEVSOVRTK	523

RESULT_39	XCPC_XENLA	STANDARD:	PRT: 1290 AA.
AC	PC50532;		
DT	01-OCT-1996 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DT	15-JUL-1998 (Rel. 36, Last annotation update)		
DE	Chromosome assembly protein XCAP-C.		
GN	XCAP-C.		
OS	Xenopus laevis (African clawed frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;		
OC	Xenopodinae; Xenopus.		
OX	NCBI_TaxID=8355;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=95042742; PubMed=7954811;		
RA	Hirano T., Michelson T.J.,		
RT	"A heterodimeric coiled-coil protein required for mitotic chromosome		
RL	condensation in vitro.";		
CC	Cell 79:449-458(1994).		
CC	-1- FUNCTION: REQUIRED FOR BOTH ASSEMBLY AND STRUCTURAL MAINTENANCE OF		
CC	MITOTIC CHROMOSOMES.		
CC	-1- SUBUNIT: ASSOCIATES WITH XCAP-E PROBABLY AS HETERODIMER.		
CC	-1- SUBCELLULAR LOCATION: NUCLEAR. REPLICATED CHROMOSOME ASSEMBLY IN		
CC	MITOTIC EXTRACTS, XCAP-C/E WAS RECRUITED TO THE CHROMATIN AND		
CC	FORMED A DISCRETE INTERNAL STRUCTURE WITHIN ASSEMBLED CHROMOSOMES.		
CC	-1- DOMAIN: CONSISTS OF TWO PUTATIVE CENTRAL COILED-COIL REGIONS		
CC	FLANKED BY PUTATIVE GLOBULAR REGIONS AT THE N- AND C-TERMINUS.		
CC	-1- SIMILARITY: BELONGS TO THE SMC FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
DR	EMBL: U13673; AAA64679.1; -		
DR	InterPro: IPR003439; ABC_transportr.		
DR	InterPro: IPR003405; SMC_C.		
DR	InterPro: IPR003395; SMC_N.		
DR	Pfam: PF02483; SMC_C; 1.		
DR	Pfam: PF02463; SMC_N; 1.		
KM	Mitosis; ATP-binding; Coiled coil; Nuclear protein.		
FT	NP_BIND 107 114		
FT	DOMAIN 264 594		
FT	DOMAIN 764 1027		
FT	DOMAIN 1094 1129		
FT	DOMAIN 1263 1290		
FT	DOMAIN 60 65		
FT	DOMAIN 747 750		
FT	DOMAIN 841 844		
FT	DOMAIN 1196 1220		
FT	SEQUENCE 1290 AA; 146988 MM; 293124924FFE90F6 CRC64;		

Query Match	13.7%;	Score 146;	DB 1;	Length 1290;
Best Local Similarity	21.2%;	Pred. No. 0.38;		





Db 1257 LSEARGKNEEIQRSISELTTOKSRLOTEAGELSRQLEEKESIVSQLSRSKQAF7QOTEEL 1316  
Oy 129 IRDMGVGOSAVEOLAVY-----CVSLKKEYENLKEARKASGEVADKLKRDLFSSRSKL 181  
Db 1317 KRQLEENKAKNALAHALQSSRHDDCLIREQYEEEOEGKAELORALSKANSEVAQWRTKY 1376  
Oy 182 QT---VYSELDOAKLEIKSAQKDLQSDAKKEIMSLKKKLTWLO 220  
Db 1377 ETDAIQRTTEELEEAKKL--AQR-LQDSEEQVEAVNAKCASTE 1416

Search completed: September 4, 2002, 16:17:35  
Job time: 494 sec